

Independent Component Analysis of Electrophysiological Data



Scott Makeig

Institute for Neural Computation
University of California San Diego

29th EEGLAB Workshop
Bois Perché, Aspet, France

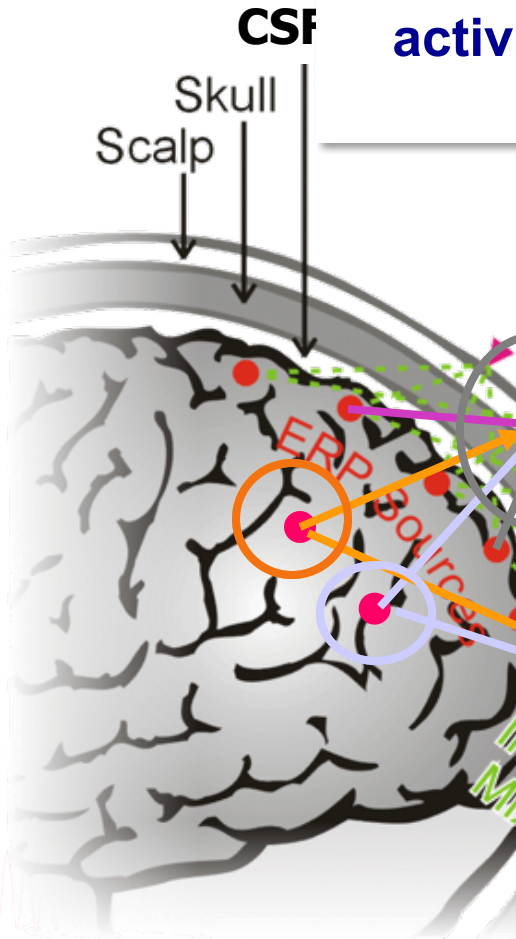
June 17-21, 2019

Blind EEG Source Separation by Independent Component Analysis



Tony Bell,
developer
of Infomax
ICA

ICA can find distinct EEG source activities -- and their 'simple' scalp maps!



Independent Component Analysis of Electroencephalographic Data

Scott Makeig
Naval Health Research Center
P.O. Box 65122
San Diego CA 92166-5122
scott@epi.lanl.gov, shre.navy.mil

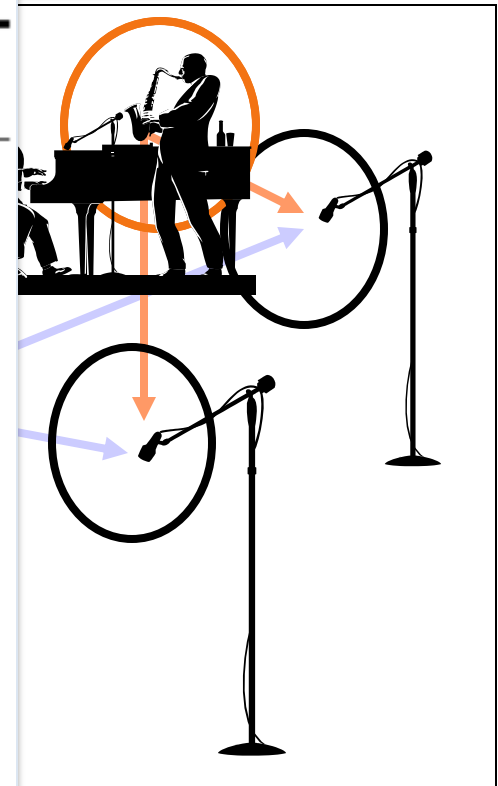
Anthony J. Bell
Computational Neurobiology Lab
The Salk Institute, P.O. Box 85800
San Diego, CA 92186-5800
tonyb@salk.edu

Tzyy-Ping Jung
Naval Health Research Center and
Computational Neurobiology Lab
The Salk Institute, P.O. Box 85800
San Diego, CA 92186-5800
jung@salk.edu

Tiziana J. Sejnowski
Howard Hughes Medical Institute and
Computational Neurobiology Lab
The Salk Institute, P.O. Box 85800
San Diego, CA 92186-5800
terry@salk.edu

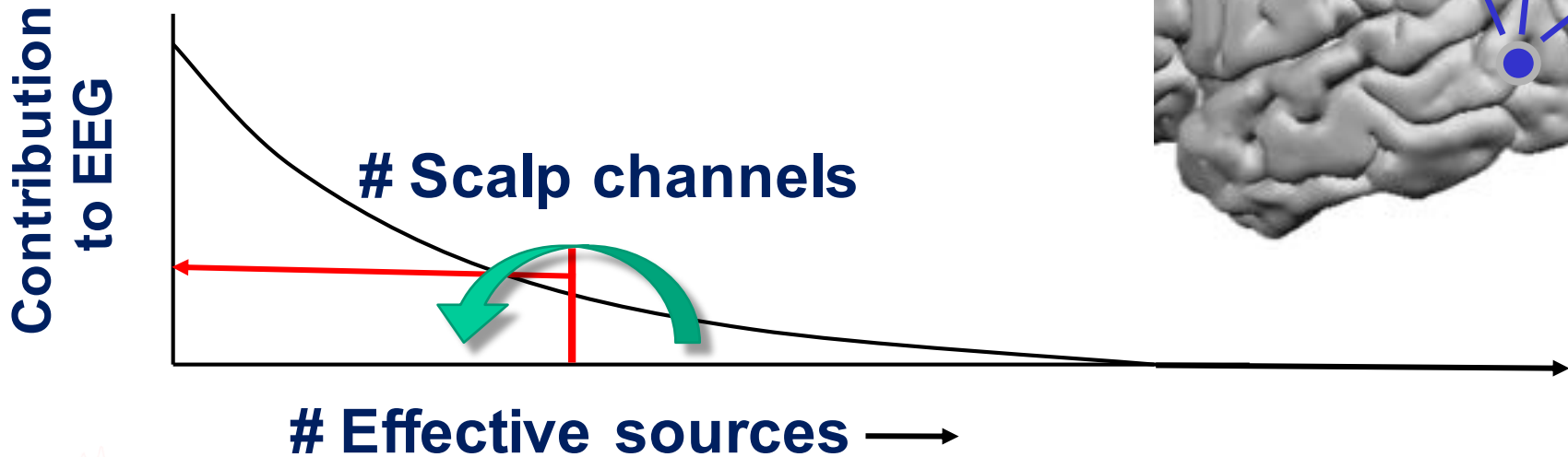
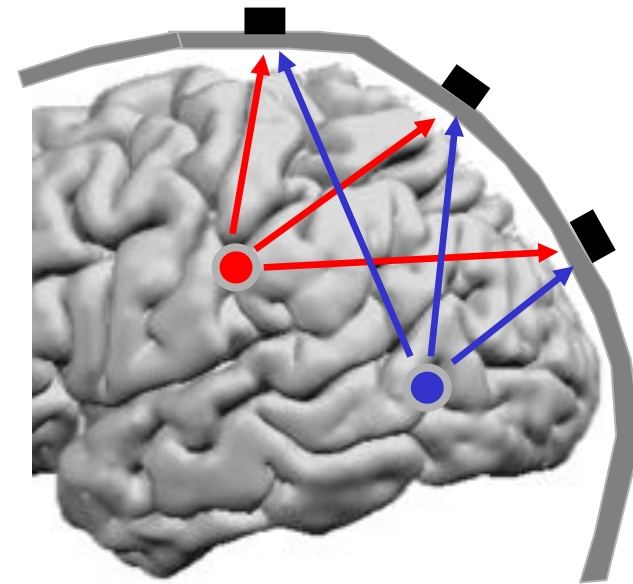
Abstract

Because of the distance between the skull and brain and their different sensitivities, electroencephalographic (EEG) data collected from any point on the human scalp includes activity generated within a large brain area. This spatial smearing of EEG data by volume conduction does not involve significant time delays, however, suggesting that the Independent Component Analysis (ICA) algorithm of Bell and Sejnowski [1] is suitable for performing blind source separation on EEG data. The ICA algorithm separates the problem of source identification from that of source localization. First results of applying the ICA algorithm to EEG and event-related potential (ERP) data collected during a sustained auditory detection task show: (1) ICA training is insensitive to different random seeds; (2) ICA may be used to segregate obvious artifactual ERP components (line and muscle noise, eye movements) from other sources; (3) ICA is capable of isolating overlapping ERP phenomena, including alpha and theta bursts and spatially-separable ERP components, to separate ICA channels; (4) Nonstationarities in EEG and behavioral state can be tracked using ICA via changes in the amount of residual correlation between ICA-filtered output channels.



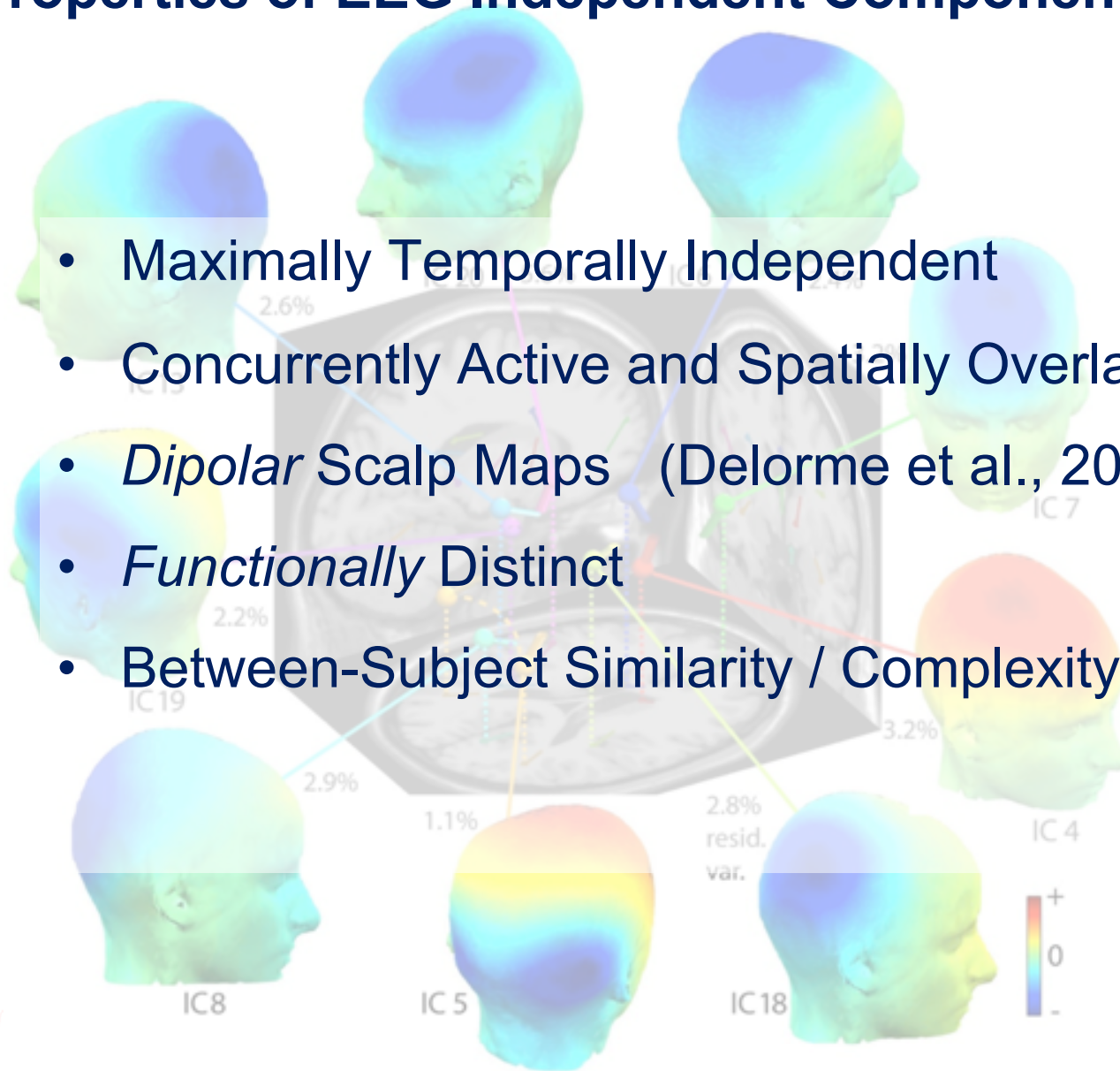
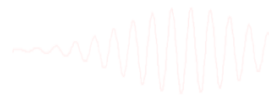
ICA Assumptions

- Mixing is **linear** at electrodes ✓
- Propagation **delays** are negligible ✓
- Component **locations** are fixed ?
- Component time courses are independent ?
- # components \leq # scalp channels ?



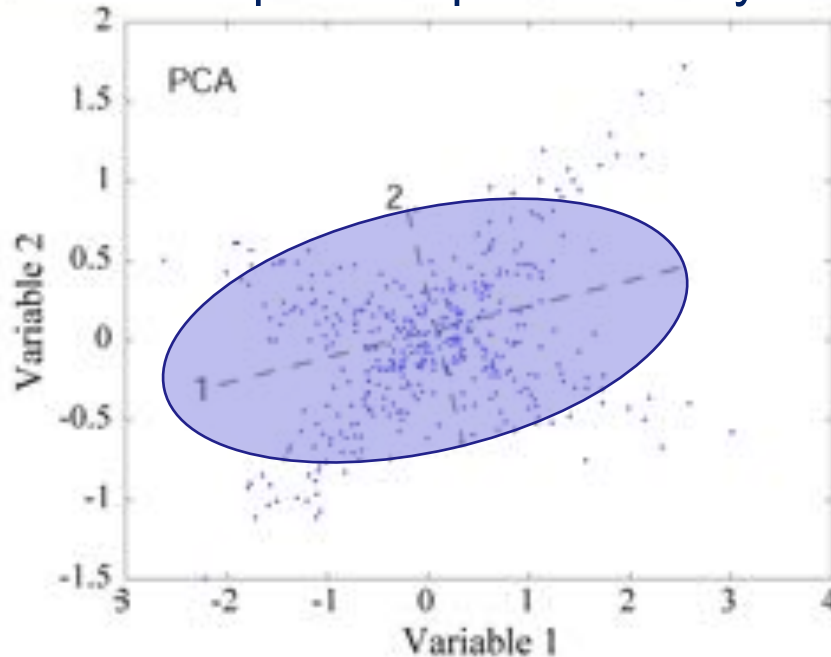
Properties of EEG Independent Components

- Maximally Temporally Independent
- Concurrently Active and Spatially Overlapping
- *Dipolar Scalp Maps* (Delorme et al., 2012)
- *Functionally Distinct*
- Between-Subject Similarity / Complexity

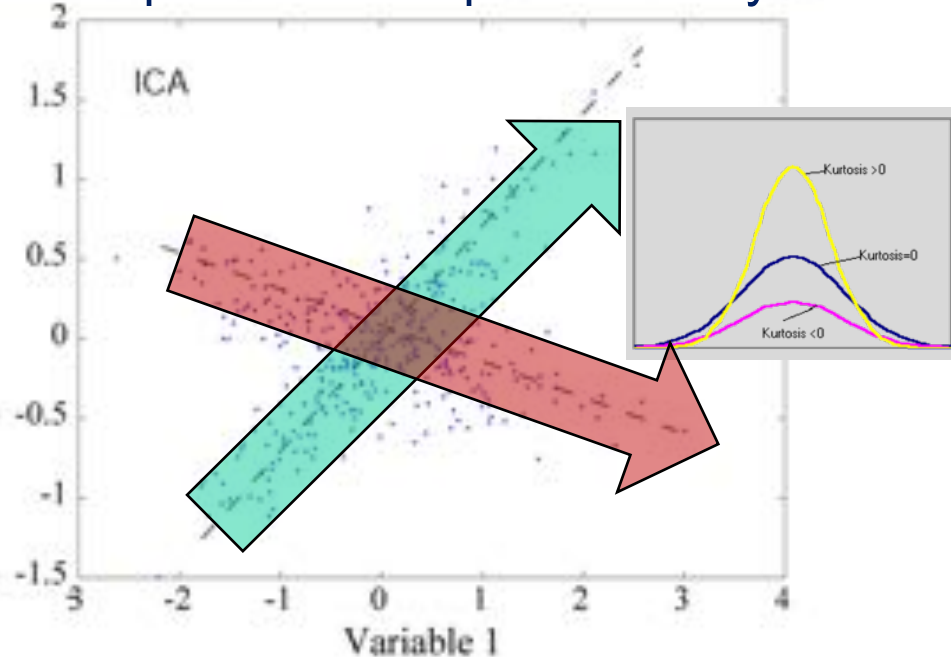


ICA vs. PCA

Principal Component Analysis



Independent Component Analysis



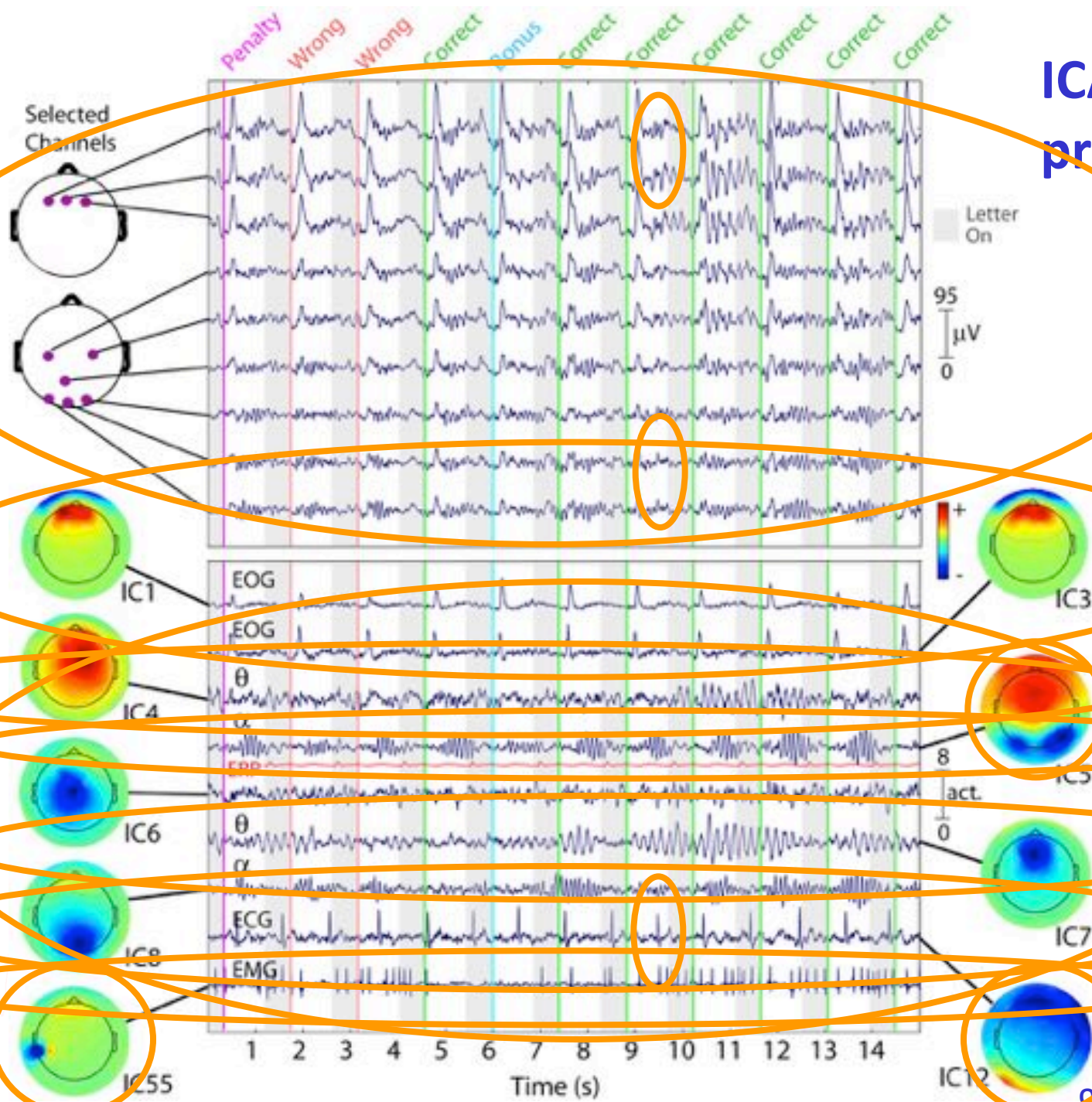
PCA simply decorrelates the outputs using an **orthogonal mixing matrix**.

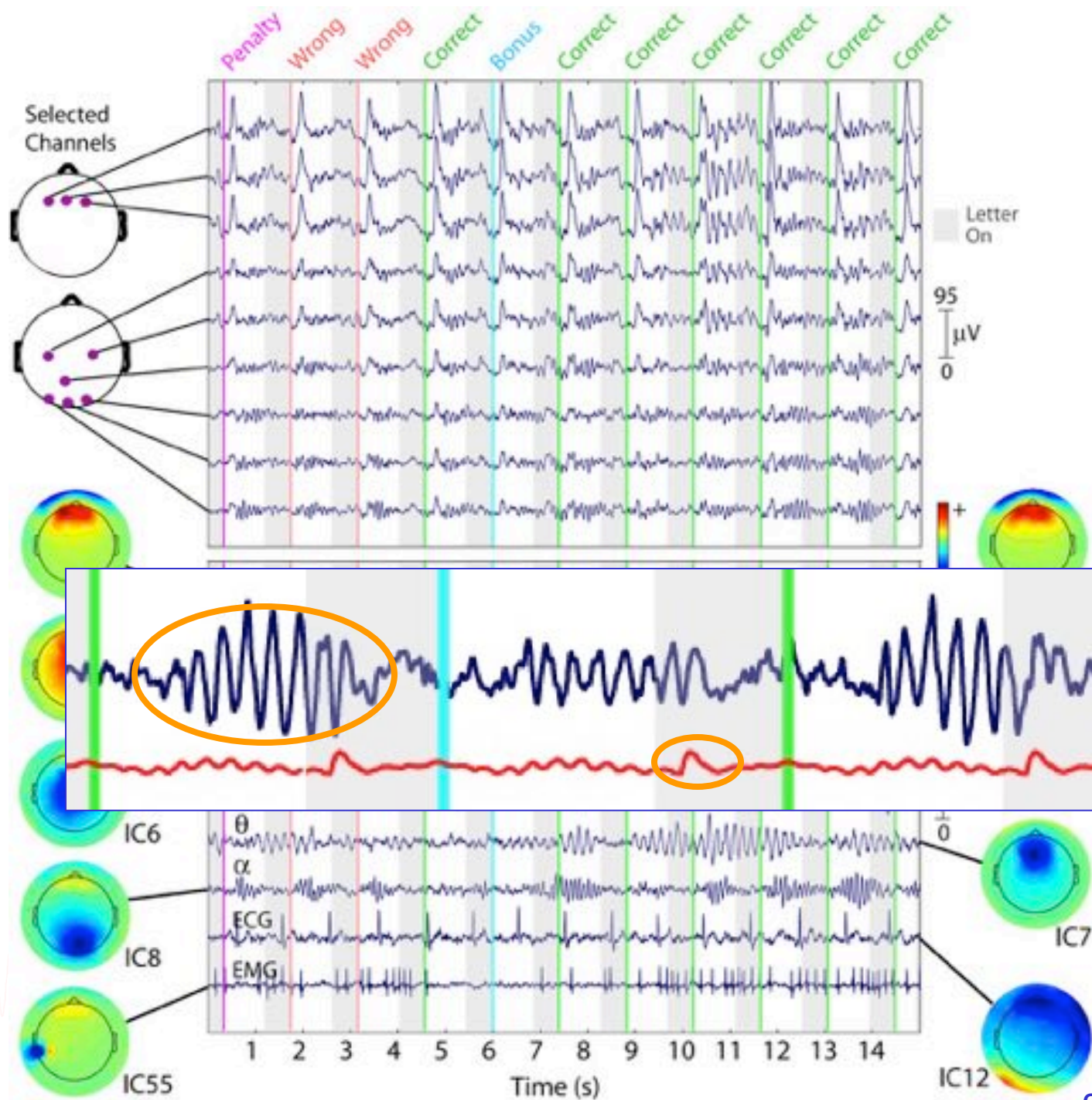
PCA makes each successive component account for as much **variance** in the data as possible.

ICA makes each component account for as much **temporally independent information** in the data as possible, with no constraints on the mixing matrix.

PCA lumps – ICA splits!

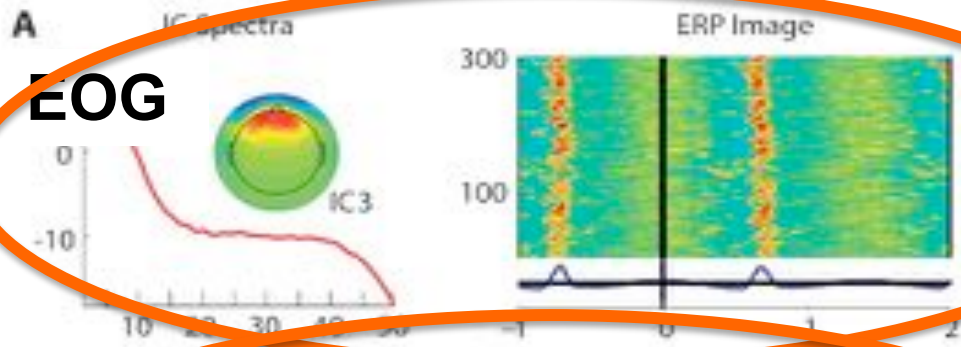
ICA in practice





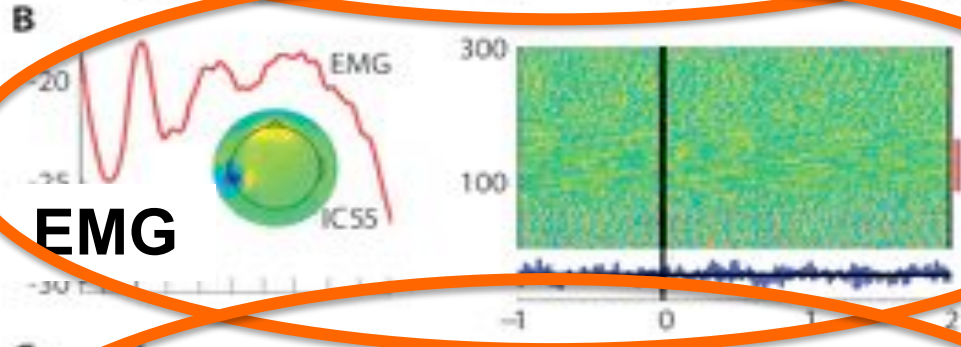
Classifying ICs

EOG



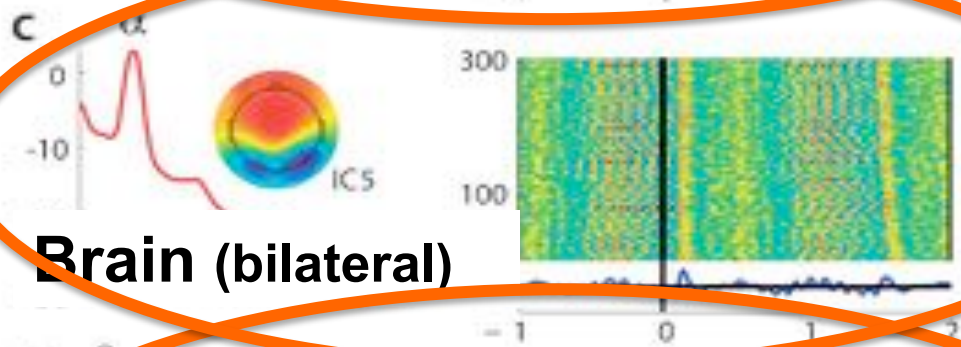
Non-brain sources

EMG



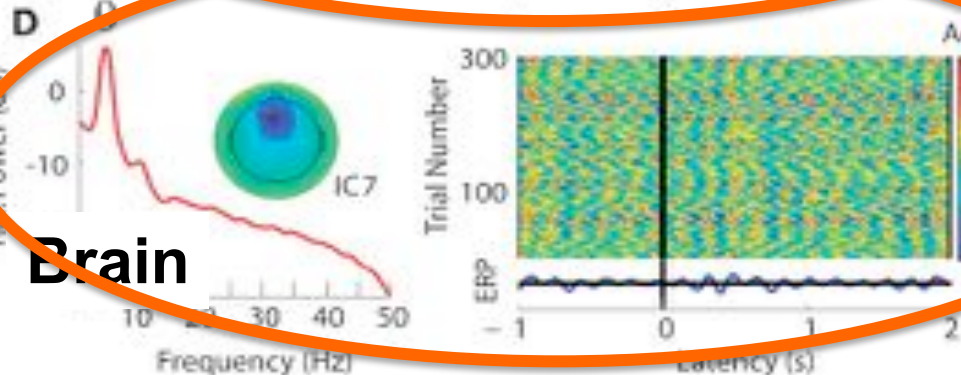
<http://ICLabel.ucsd.edu>

Brain (bilateral)

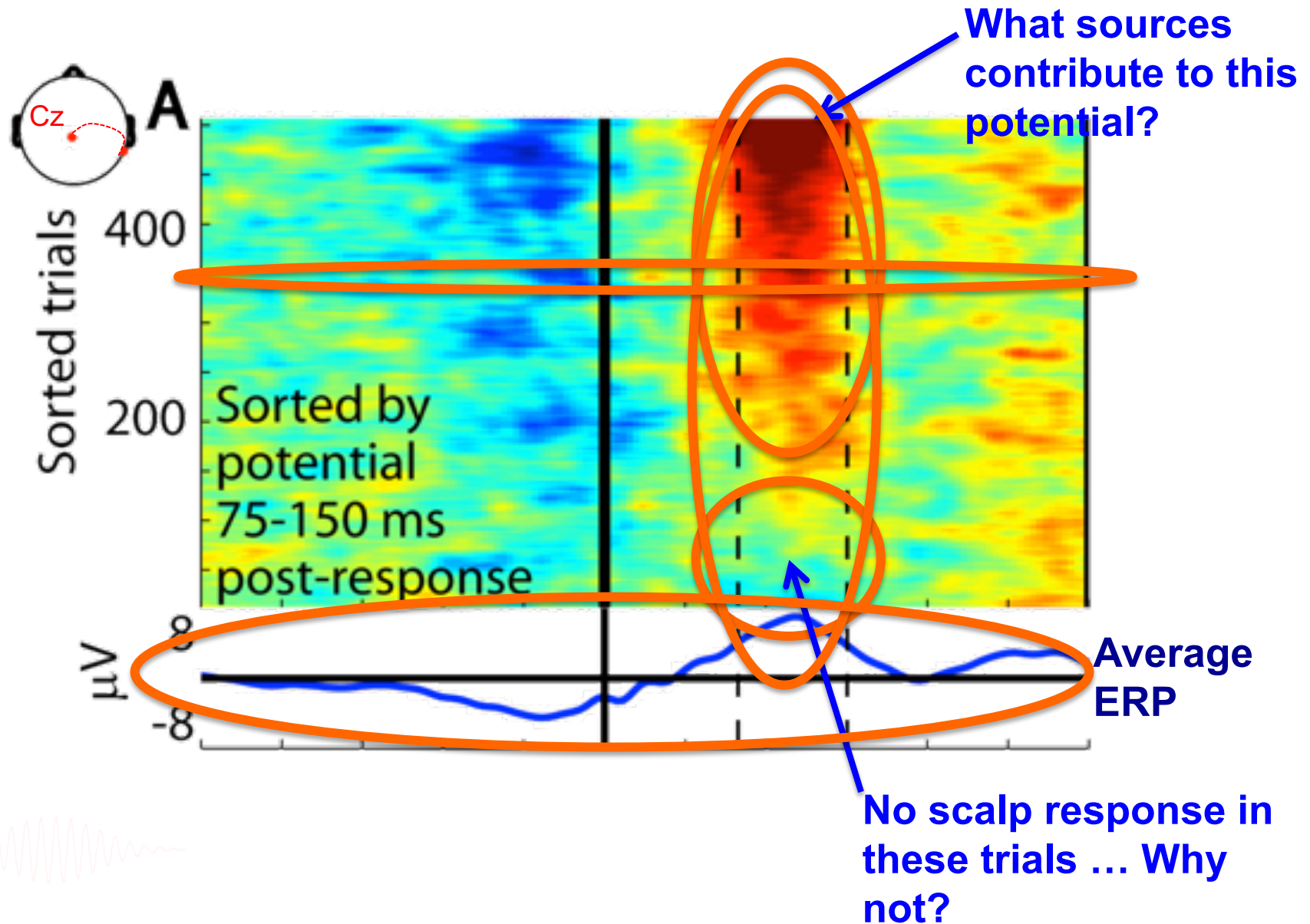


Effective brain sources

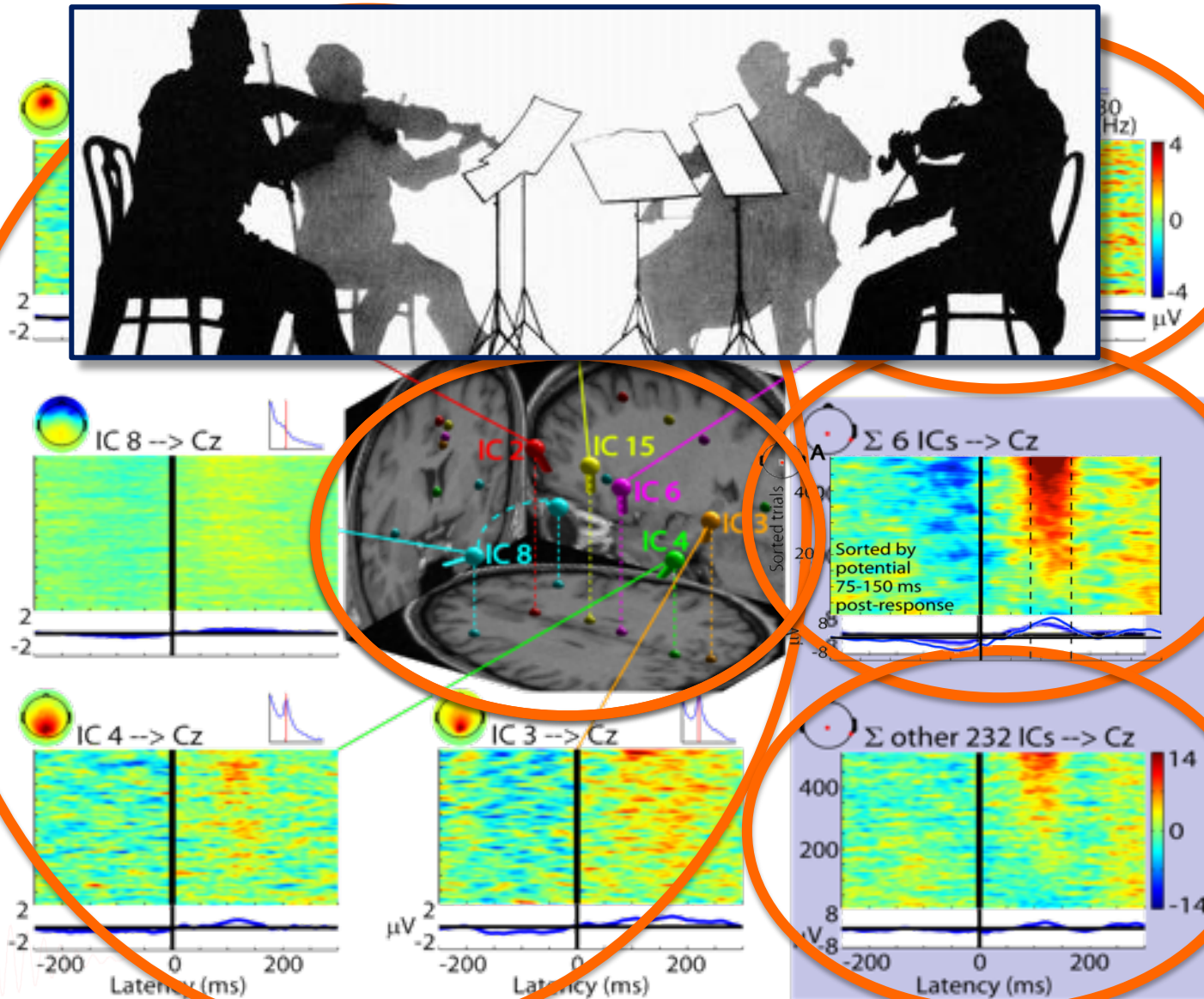
Brain



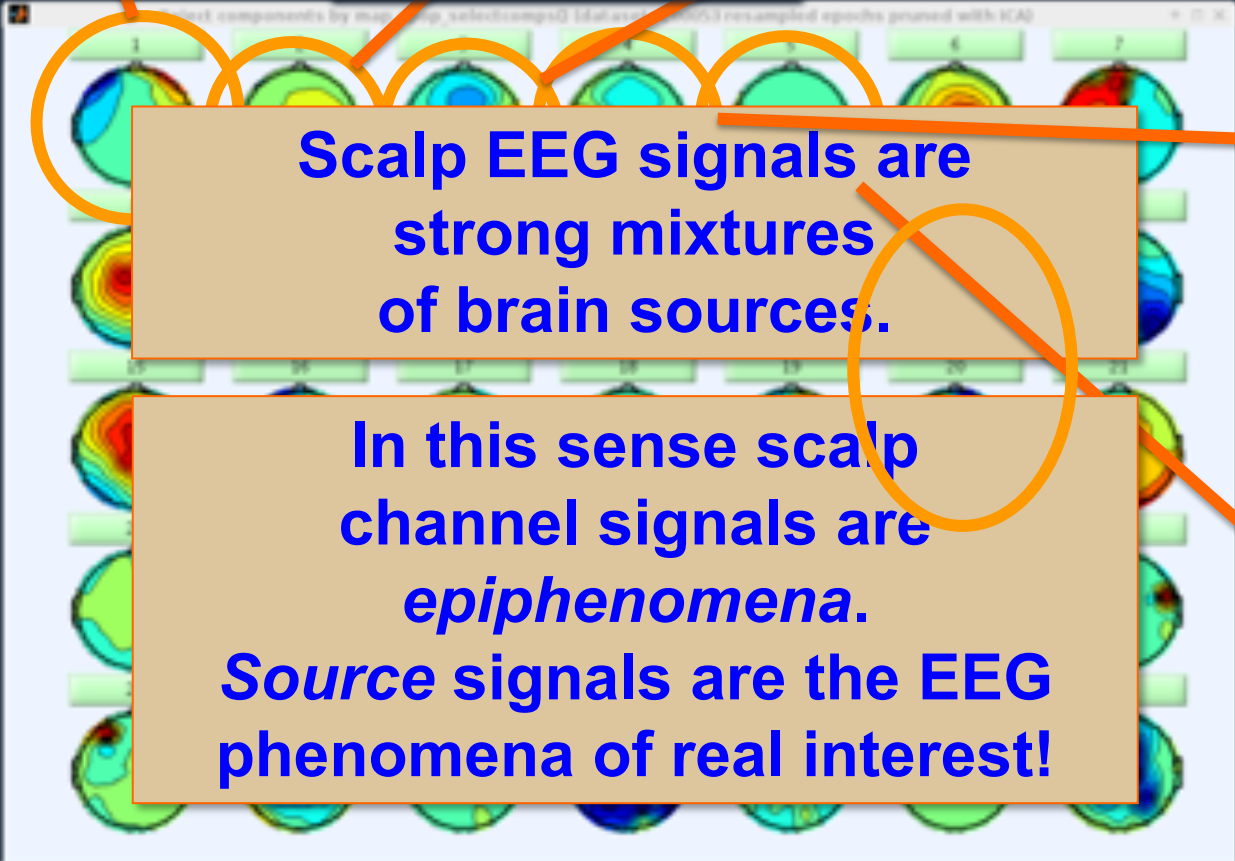
A P300' visual target response at a vertex channel



The vertex channel response sums 238 independent sources.



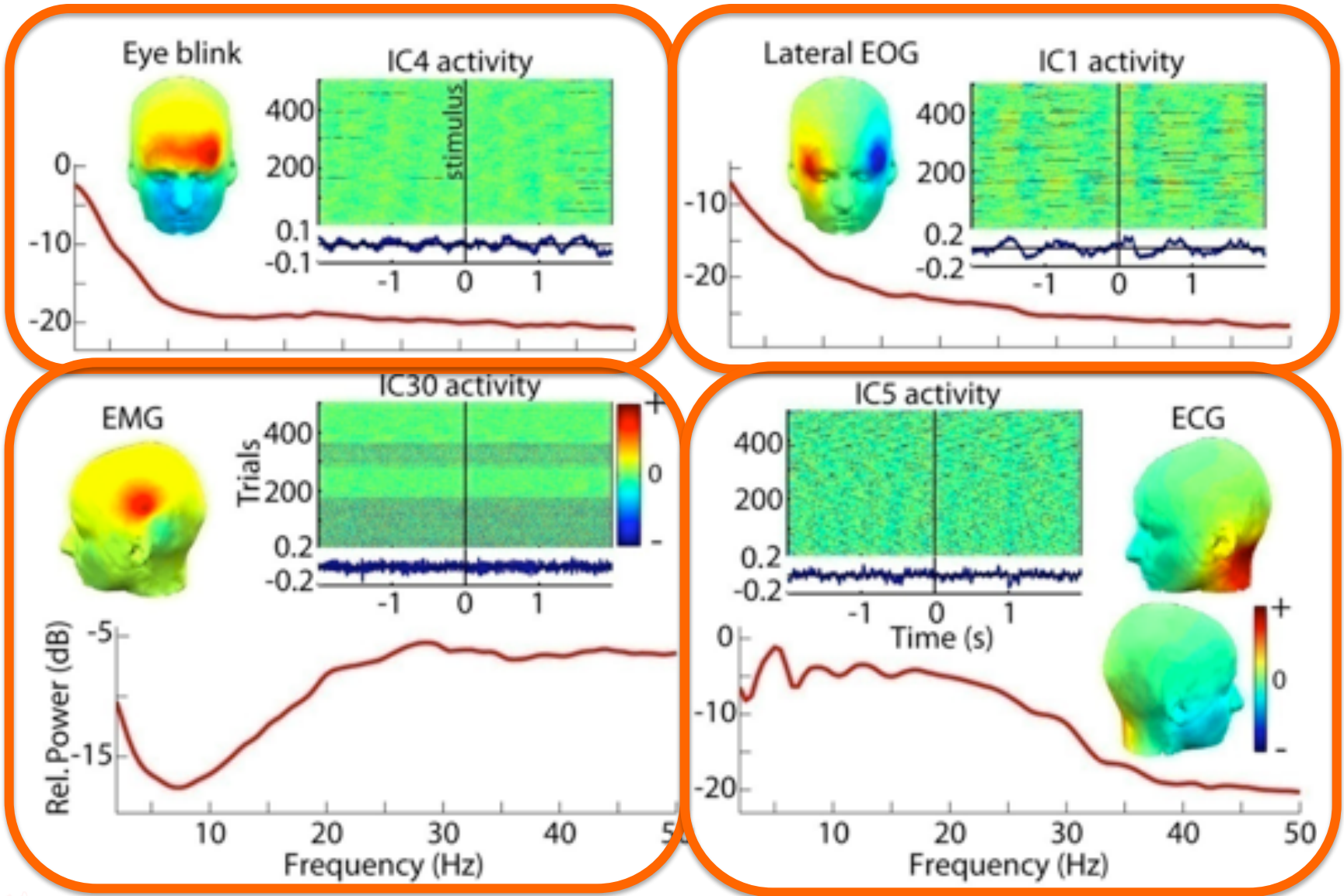
No more than
~30% of any
scalp channel
variance is
produced by any
one brain source!



Scalp EEG signals are
strong mixtures
of brain sources.

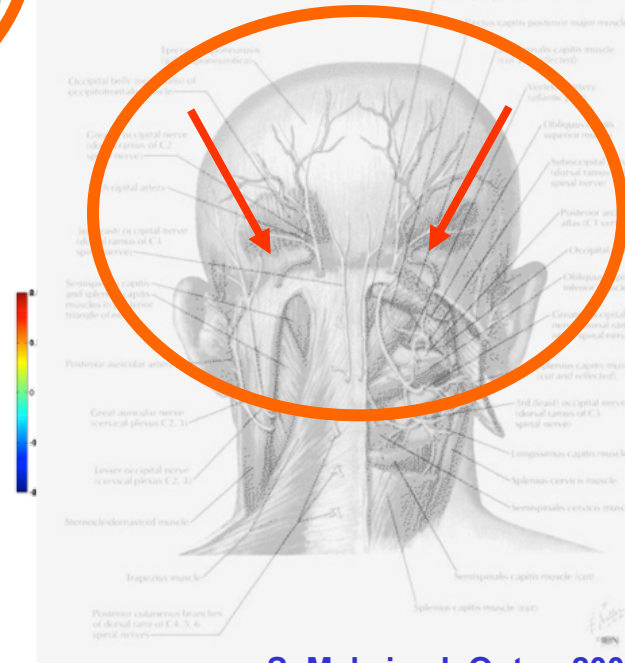
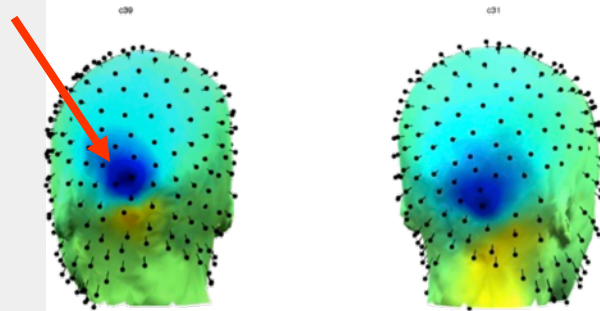
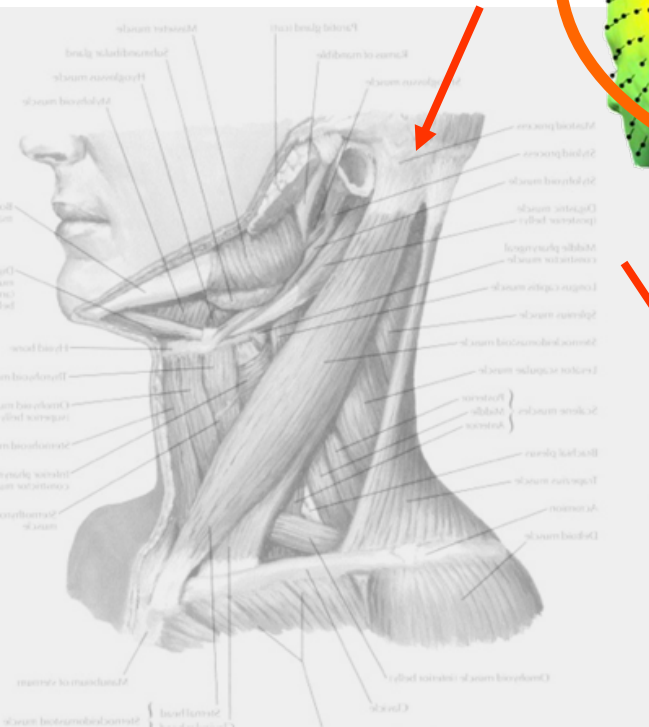
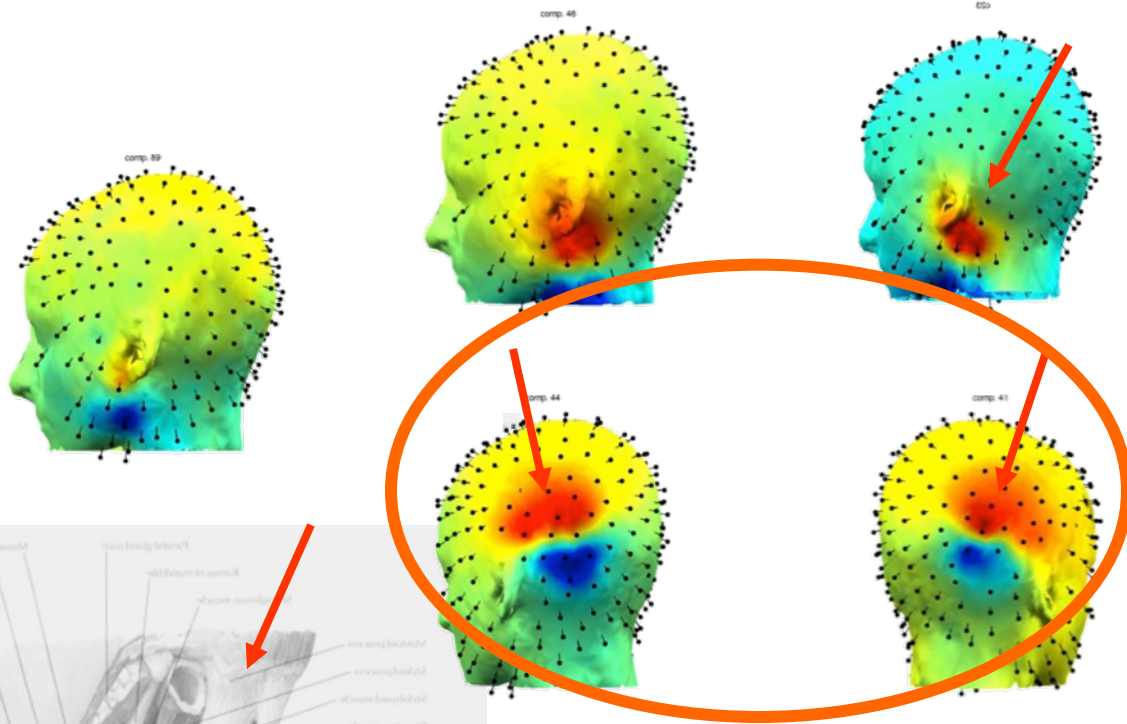
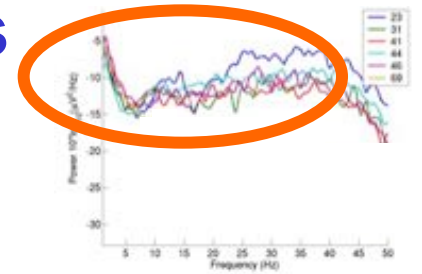
In this sense scalp
channel signals are
epiphenomena.
Source signals are the EEG
phenomena of real interest!

ICA finds Non-Brain Independent Component (IC) Processes ...

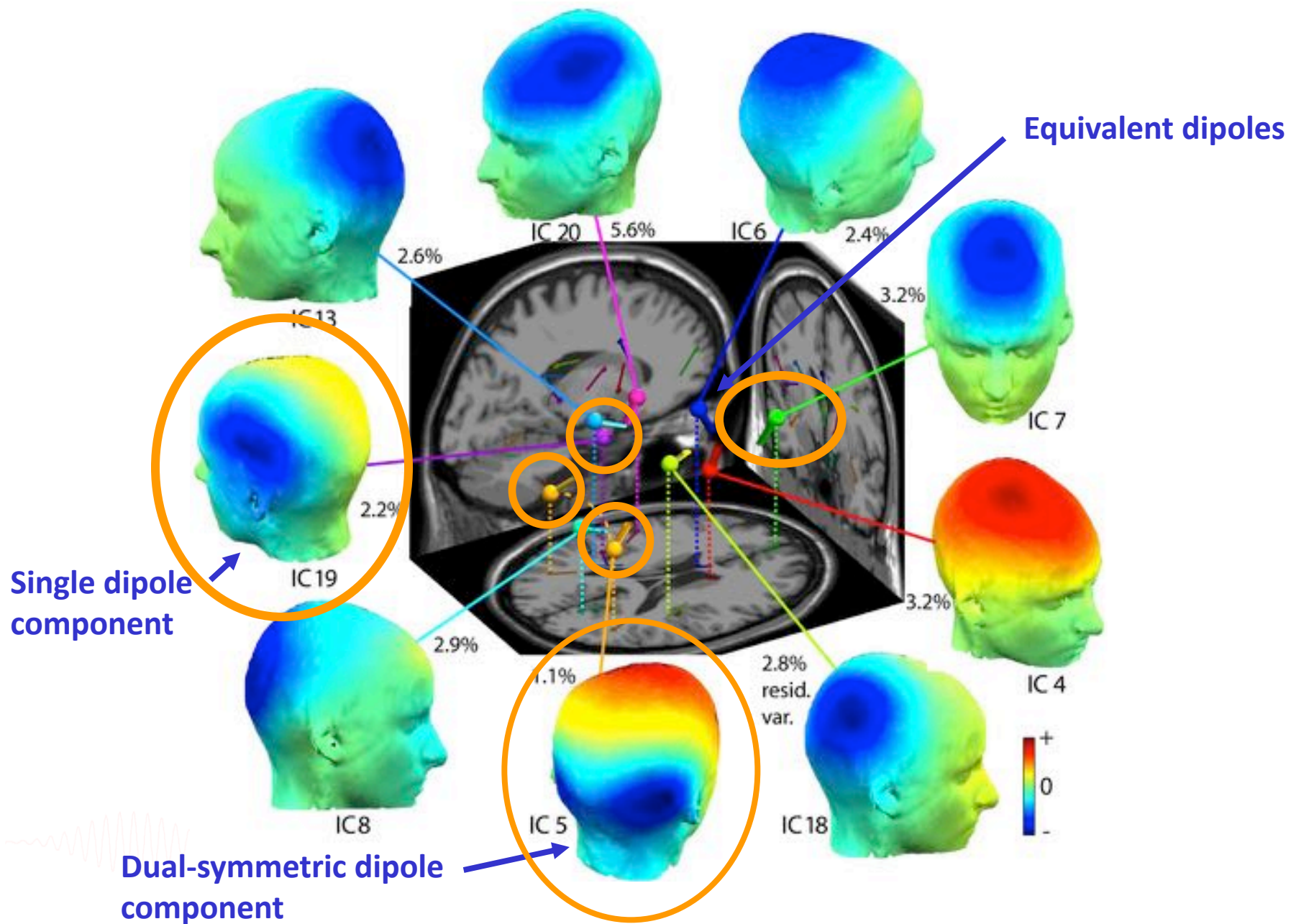


... separates them from the remainder of the data ...

Independent muscle signals



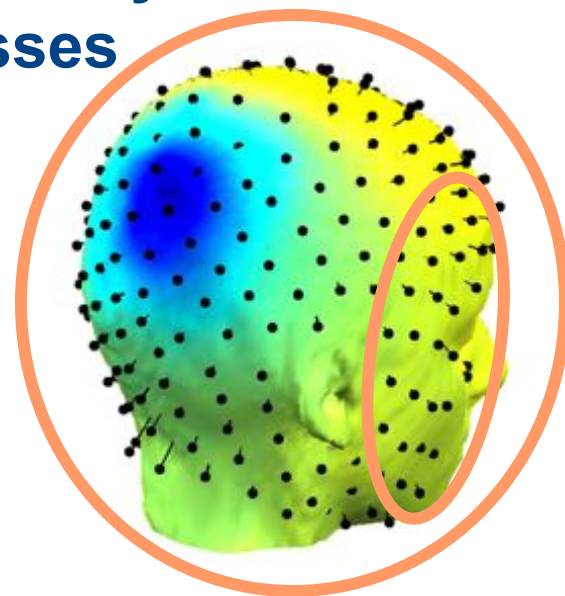
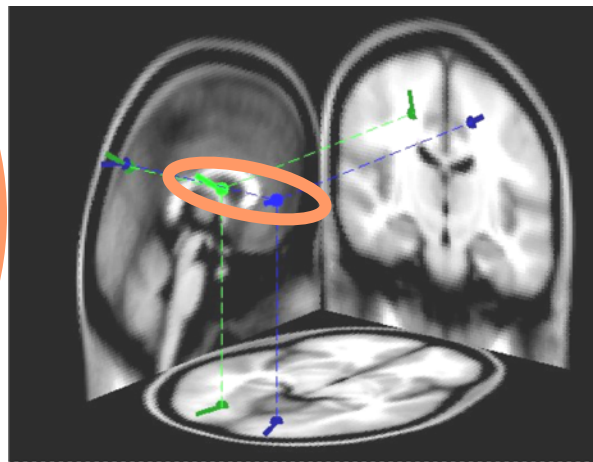
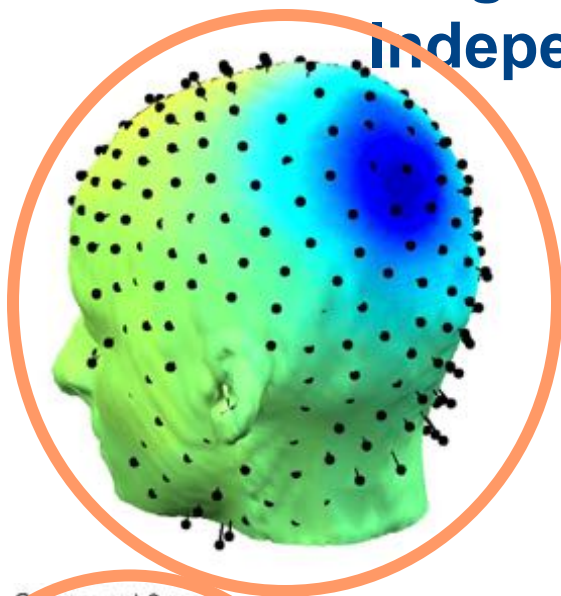
... and also separates cortical brain IC processes



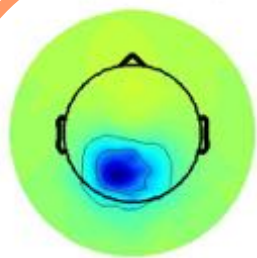
IC9

Single Session - Two Maximally Independent Alpha Processes

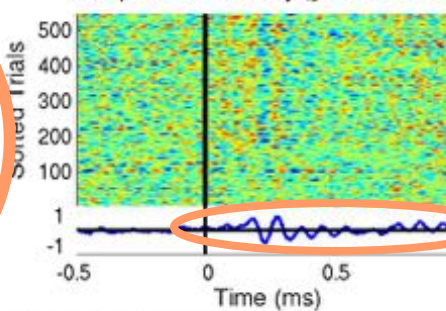
IC11



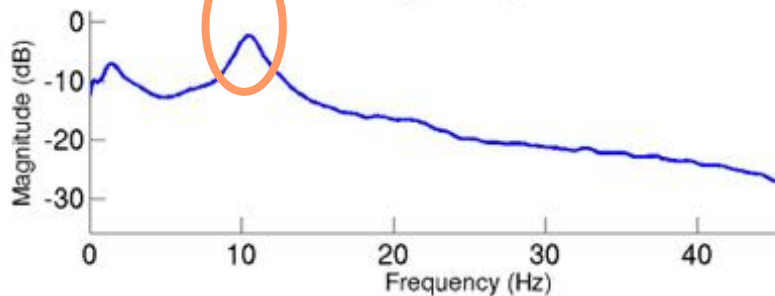
Component 9 map



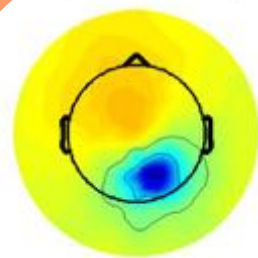
Component 9 activity (global offset 0.02)



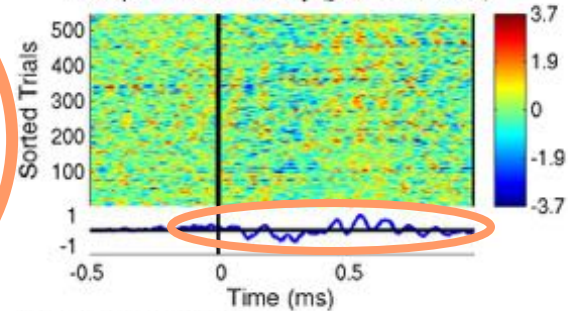
Activity power spectrum



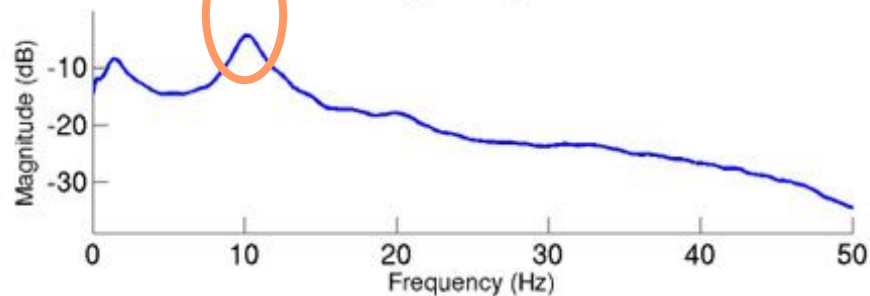
Component 11 map



Component 11 activity (global offset -0.038)

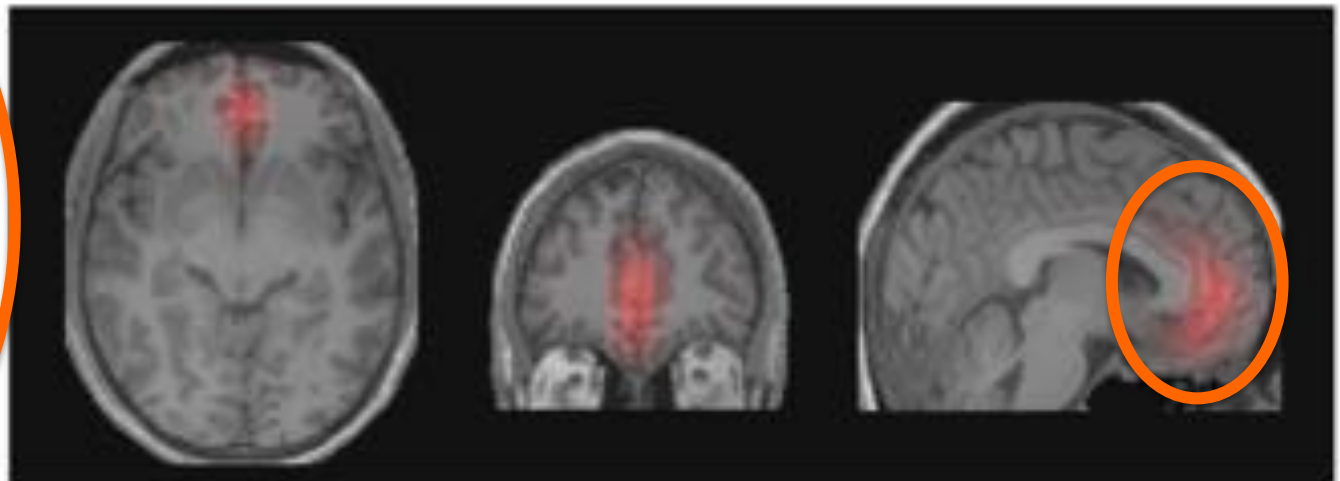
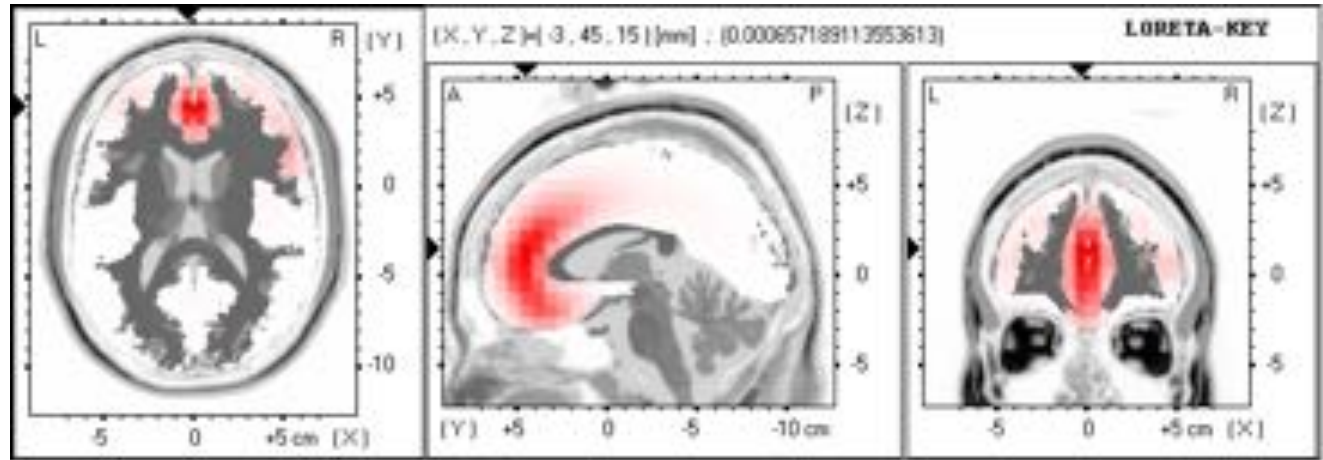
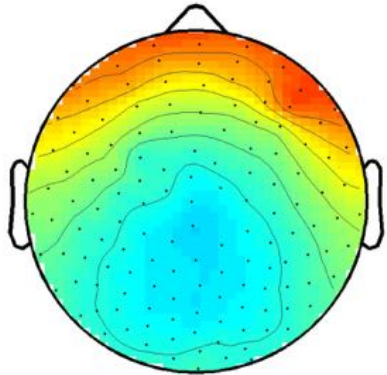


Activity power spectrum



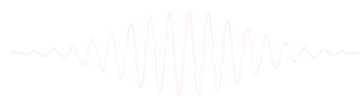
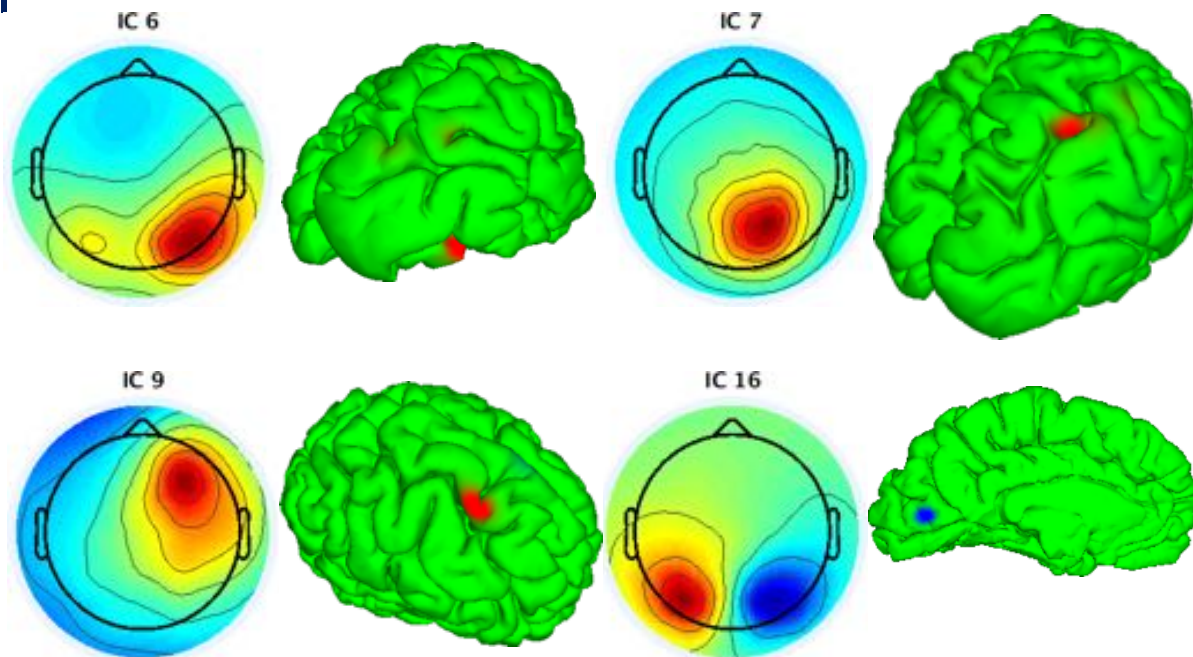
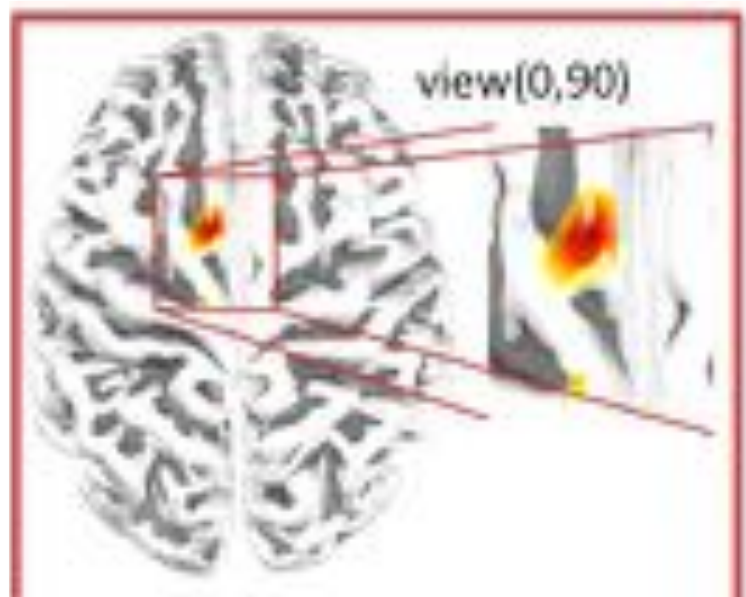
EEG Source Localization

LORETA = Low-Resolution Electrical Tomography

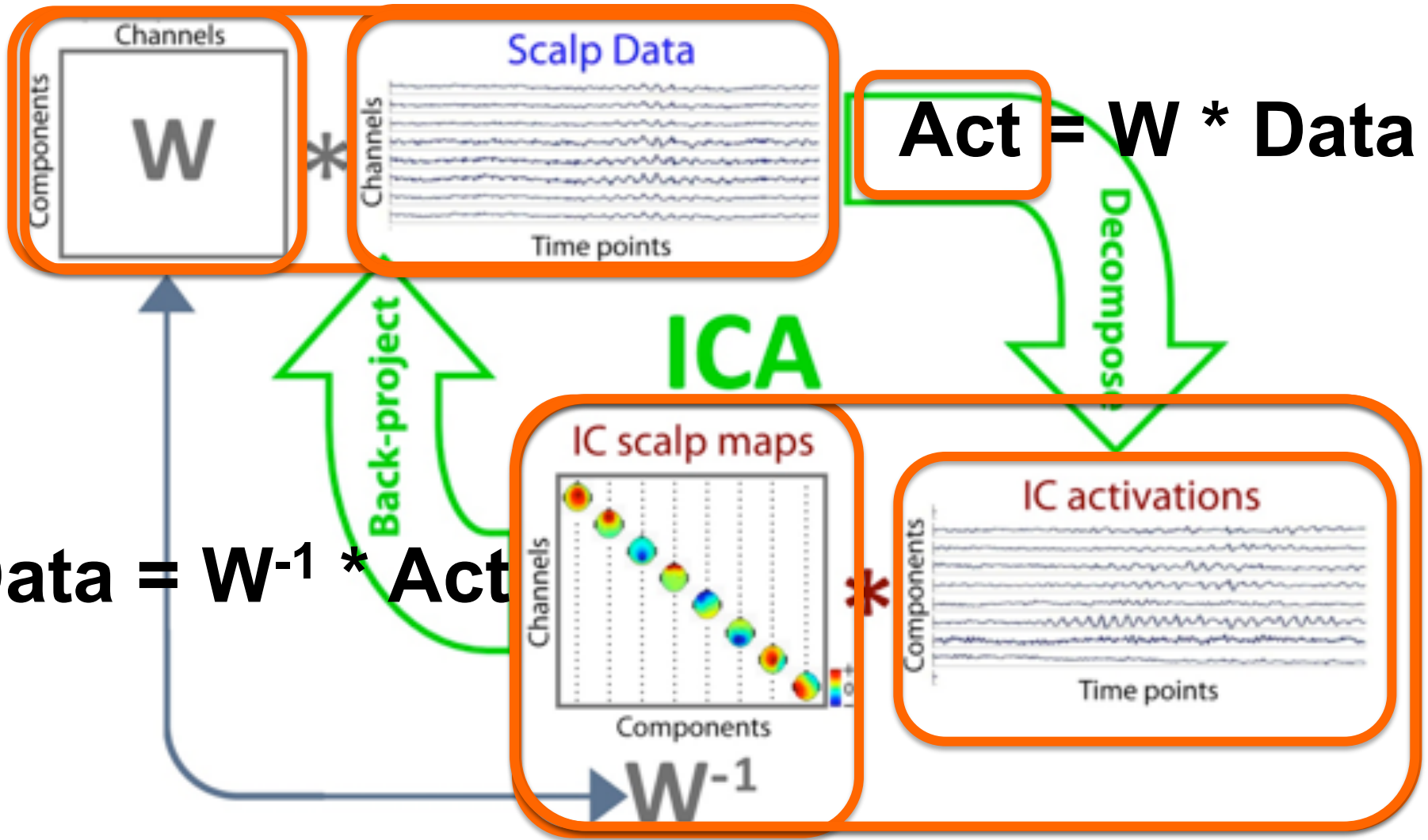


Sparse
Compact
Smooth
(SCS)

Distributed IC source
location estimation
using SCALE-optimal
head model.



ICA is a linear data decomposition method



Infomax ICA learning approach

How to make the outputs statistical independent?

Minimize their redundancy or mutual information.

Consider the joint entropy of two components,

$$H(y_1, y_2) = H(y_1) + H(y_2) - I(y_1, y_2).$$

Maximizing $H(y_1, y_2)$

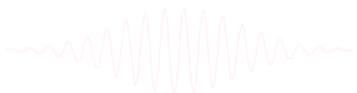
minimizing $I(y_1, y_2)$.

Infomax

The learning rule:

$$\Delta \mathbf{W} \propto \frac{\partial H(\mathbf{y})}{\partial \mathbf{W}} \underbrace{\mathbf{W}^T \mathbf{W}}_{\text{Natural gradient normalization (Amari)}}$$

Is 0 if the two variables are independent

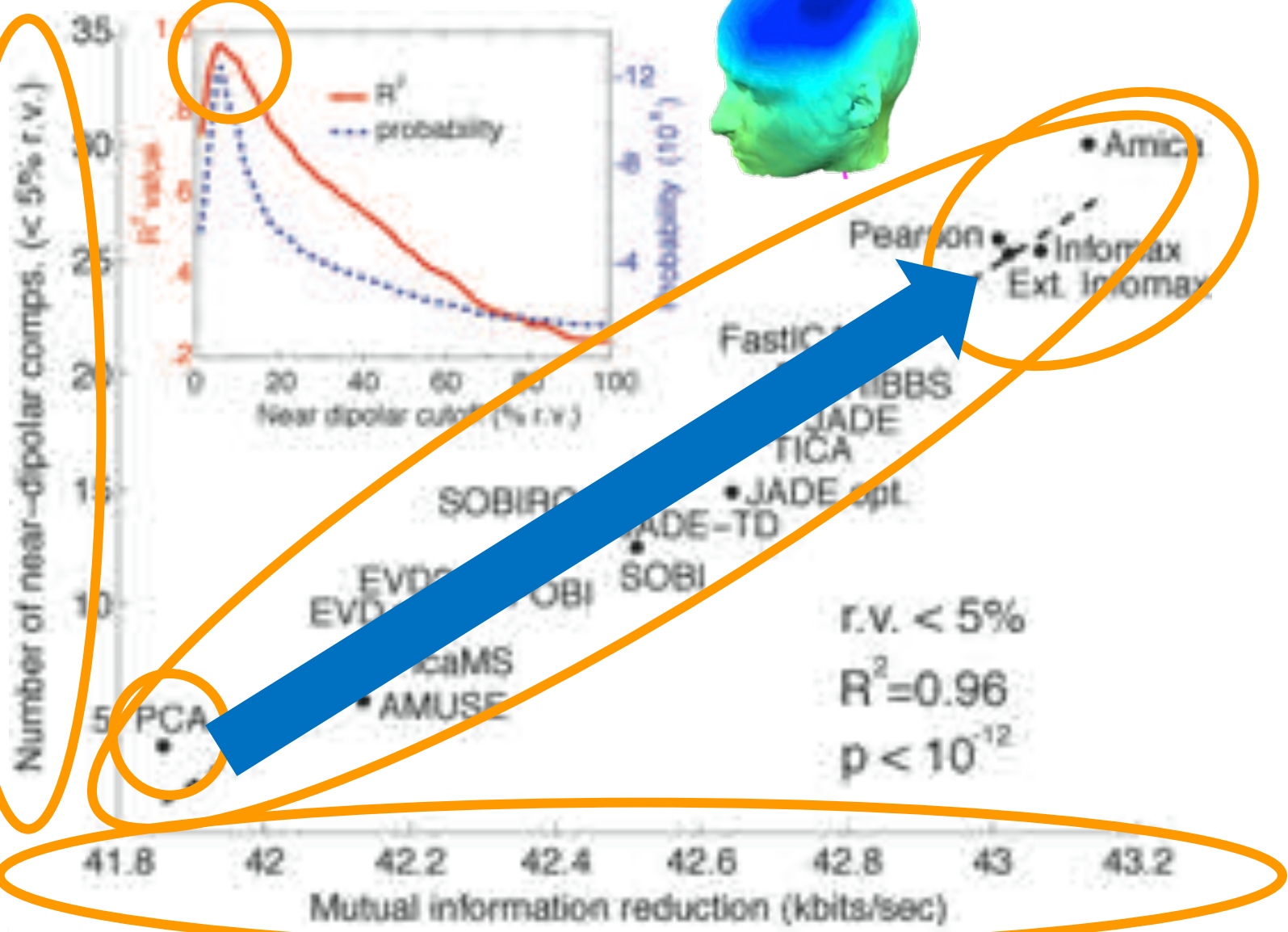


Historical Remarks

- Herault & Jutten ("Space or time adaptive signal processing by neural network models", *Neural Nets for Computing Meeting*, Snowbird, Utah, 1986): **Seminal paper**
- Bell & Sejnowski (1995): Information maximization (**Infomax**)
- Makeig, Bell, Jung, Sejnowski (1996); ICA decomposition of EEG
- Amari et al. (1996): Natural gradient learning
- Cardoso (1996): Joint approximate diagonalization (JADE)
- Hyvarinen (1999): (fastICA)
- Lee/Girolami (1999): Mixture model ICA (**Extended Infomax**)
- Palmer (2006-9): Adaptive mixture ICA (**AMICA**)

Applications of ICA to biomedical signals

- EEG/ERP analysis (Makeig, Bell, Jung & Sejnowski, **NIPS 1996**)
- fMRI analysis (McKeown et al., 1998)
- Fetal/mother ECG separation (Cardoso, 1998)
- Electrocorticography (ECoG) (Whitmer, 2010)



100%

103%

Delorme et al., *PLOS One*, 2012

S. Makeig, 2011

Important Recent Result (2012)

Those linear decompositions of multi-channel EEG data that find ICs whose time courses are **more** temporally **independent** ...

Also find more ICs whose scalp maps are highly **'dipolar'** – i.e., ICs compatible with the spatial projection of a single local cortical (or non-brain, 'artifactual') source process – whose location and distribution can be more or less accurately estimated.

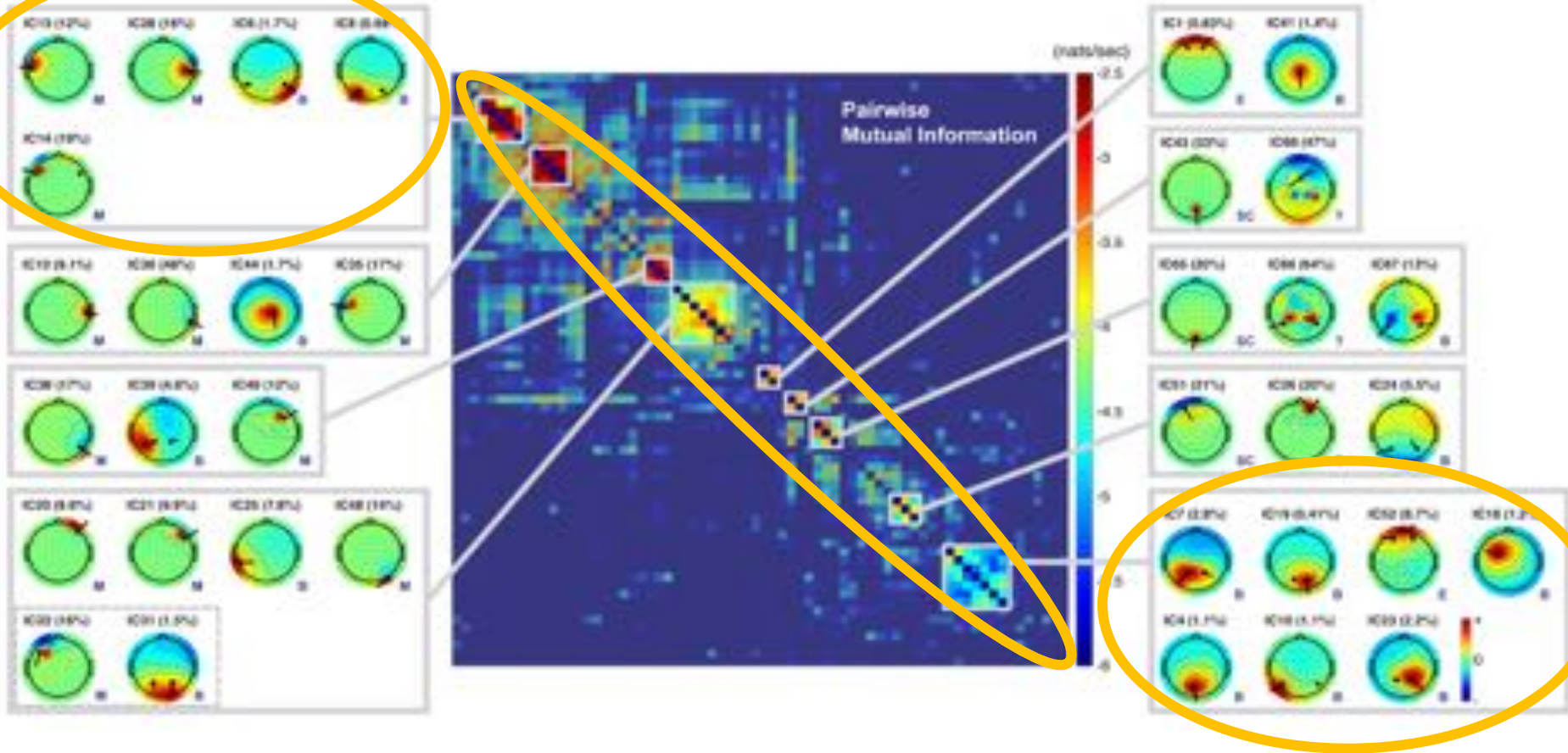
The more independent the component time courses \leftrightarrow

The larger the number of ~dipolar component scalp maps.

Hypothesis: Dipolar ICs = Localized cortical source processes

Pairwise Mutual Information Subspaces

pmi ()

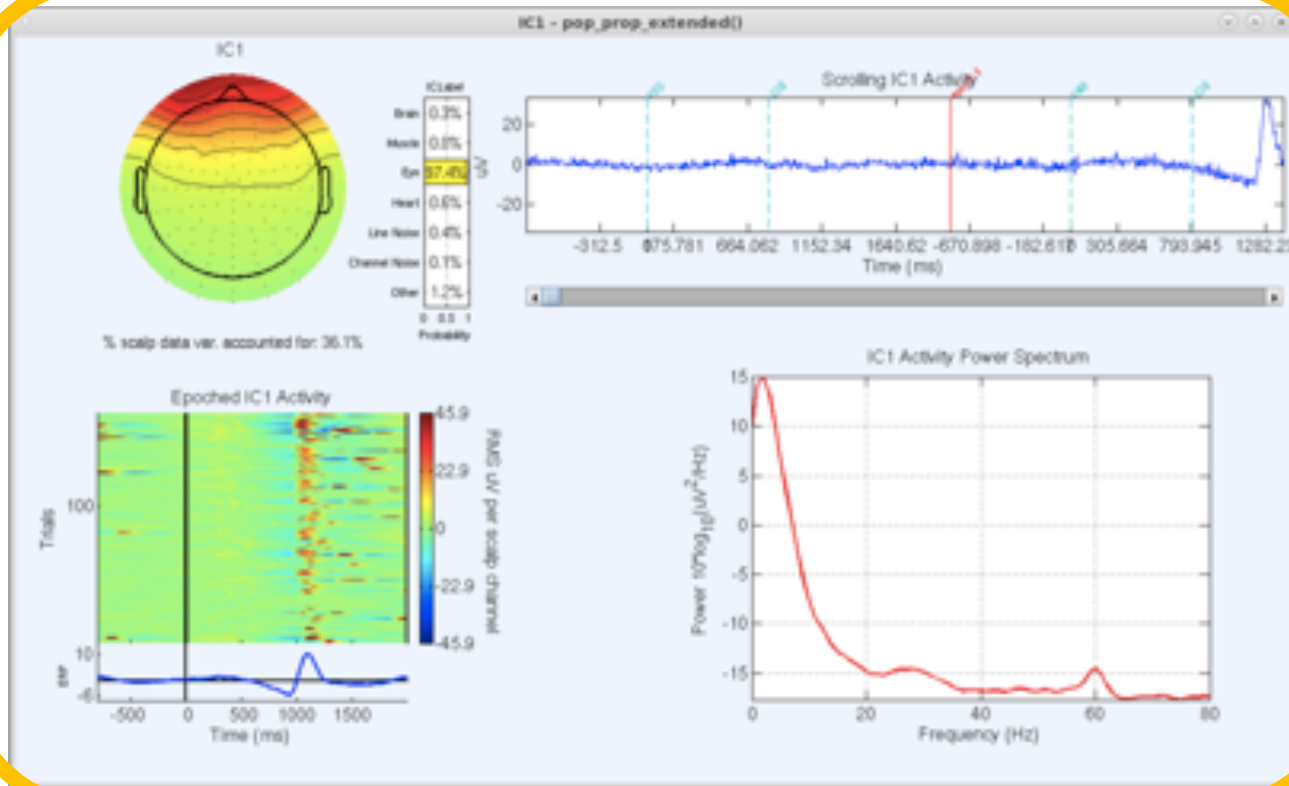


ICA transforms multiple correlated channels into maximally mutually independent component subspaces

Distinguishing IC types: ICLabel

→ <https://labeling.ucsd.edu>

Info: <https://sccn.ucsd.edu/wiki/ICLabel>



The ICLabel plug-in

Multi-model ICA: AMICA

Sleep stages:

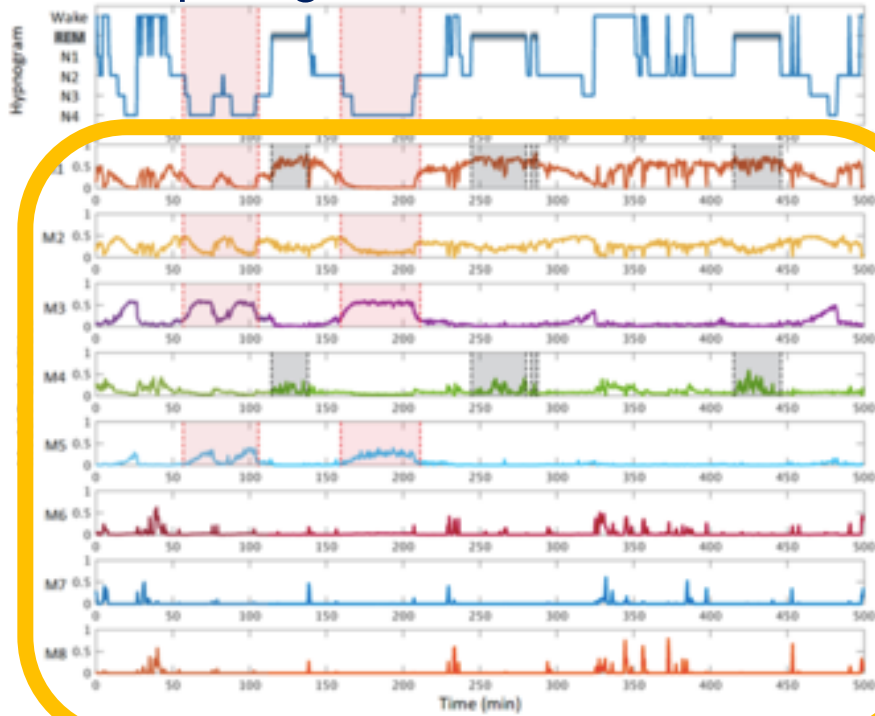


Figure 4: The top panel shows the hypnogram, i.e., sleep stages annotated from the EEG record by a sleep expert, of a sleep session from a single subject. Bottom panels show mean probabilities, within each 30-sec sleep scoring interval, of ICA models learned by an 8-model AMICA decomposition applied to the EEG record. Red-shaded regions highlight changes in model probabilities for relevant models during transitions to and periods of deep sleep (N4). Gray-shaded regions highlight probability value changes for relevant models during REM sleep.

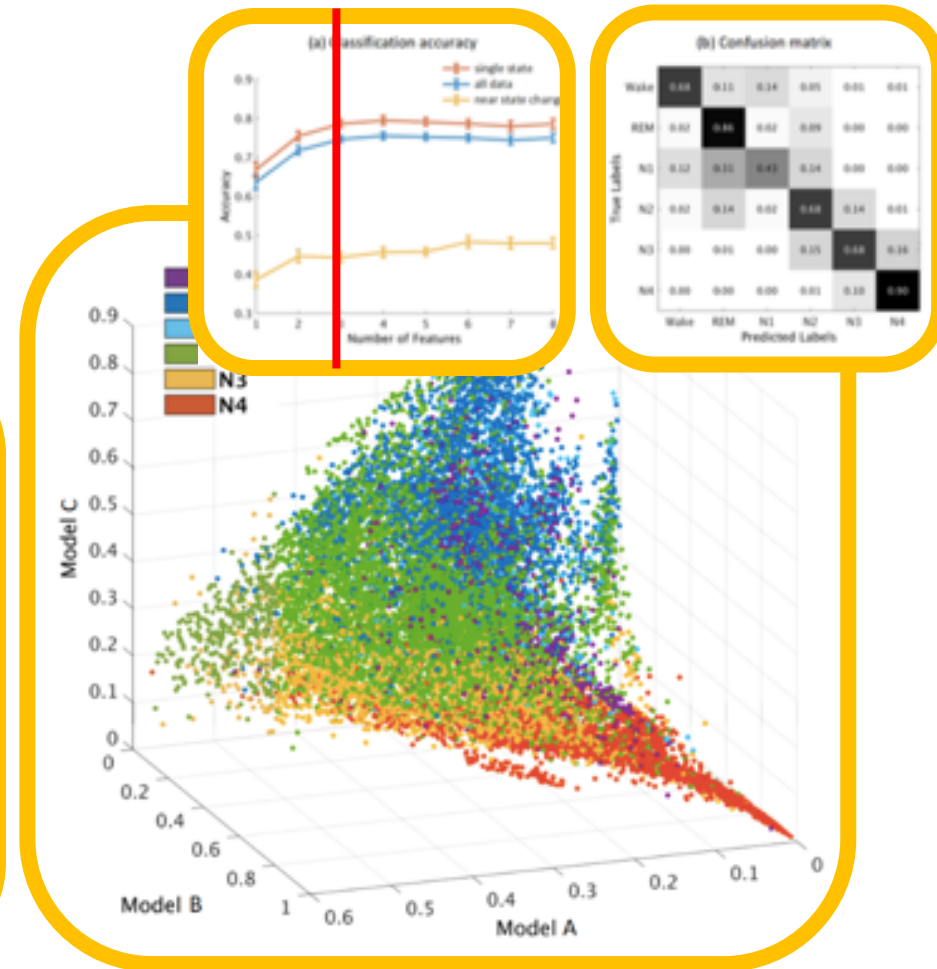
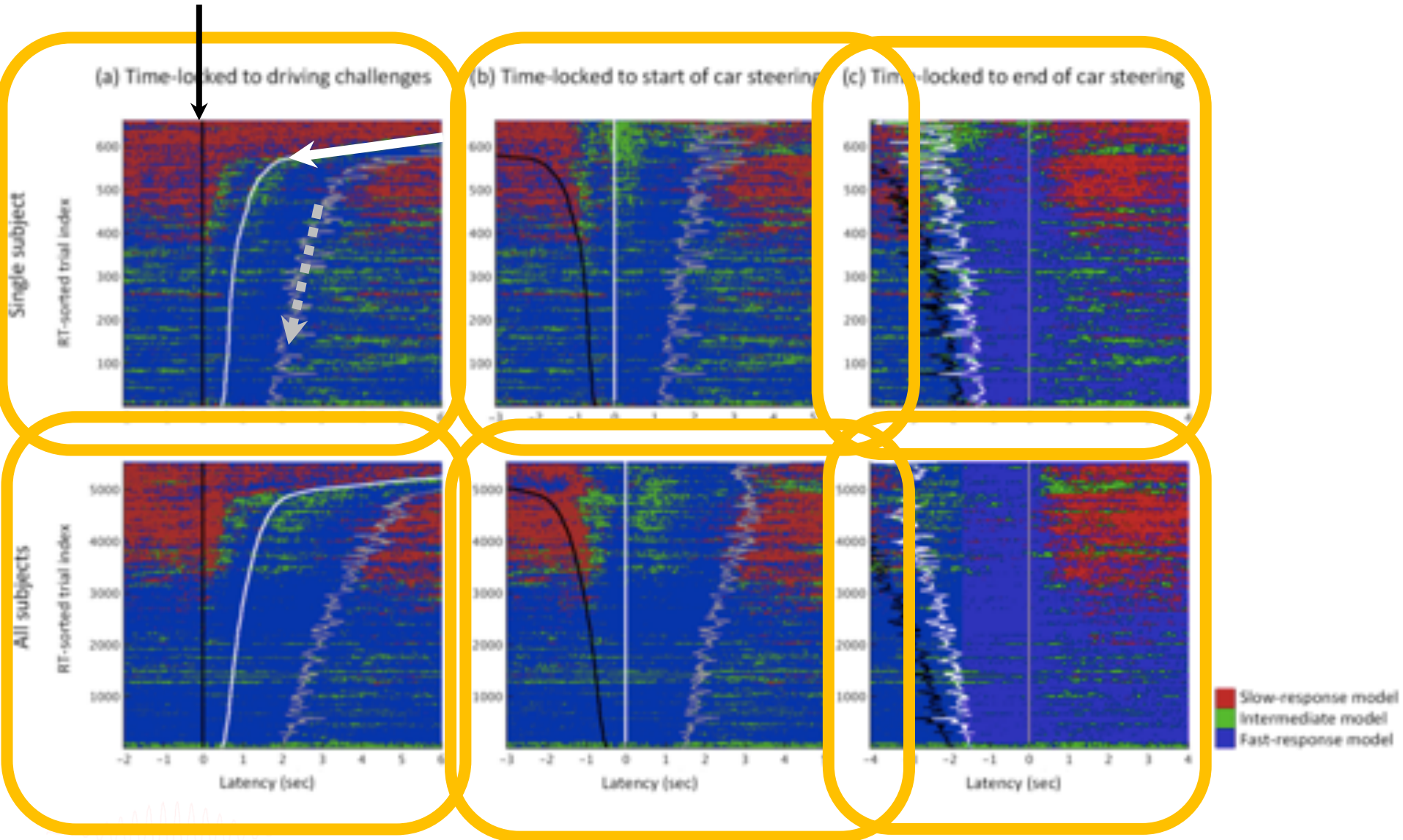
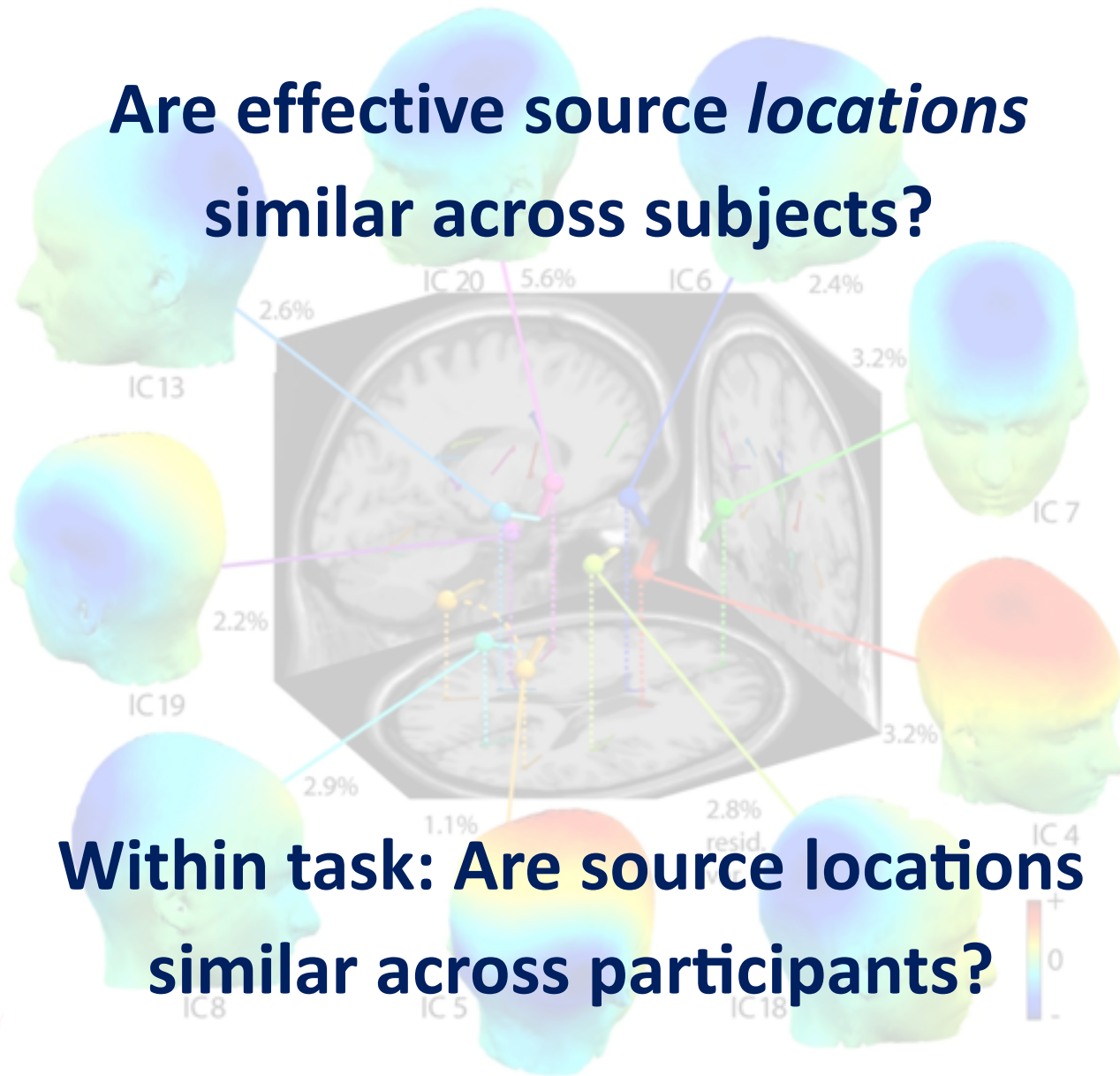


Figure 6: Scatter plot of window-mean model probabilities for AMICA model clusters A, B, and C (cf. Fig. 5), each point representing mean model probability within a 30-sec data segment from sleep recordings of 7 healthy subjects and 10 patients. Colors represent expert designated sleep-stage labels for the same data segments. Note the distinct deep sleep (N4) pattern and the relative closeness of wake and REM sleep characteristics.

Multi-model ICA: AMICA



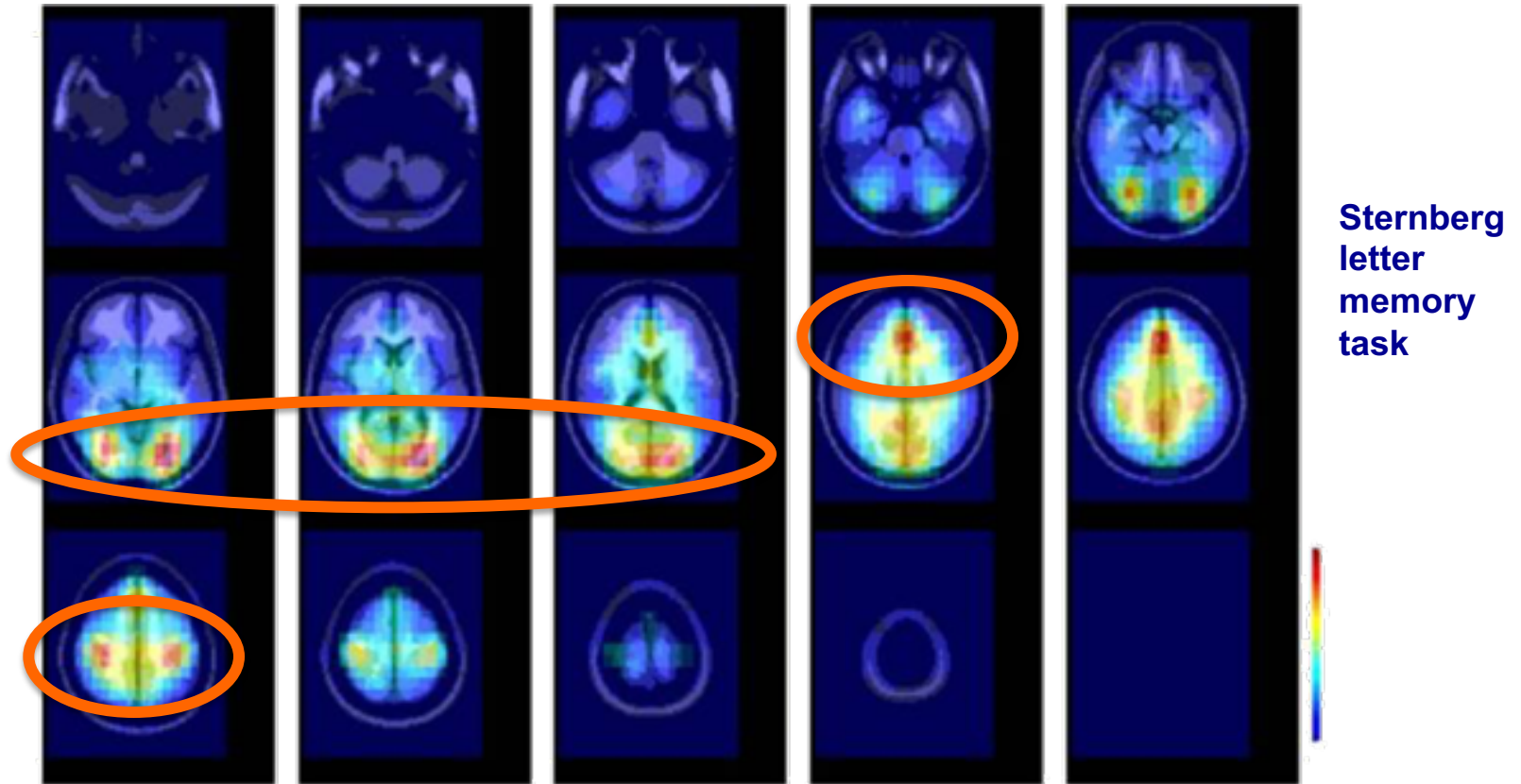
Are effective source *locations* similar across subjects?



Within task: Are source locations similar across participants?

Effective Source Density

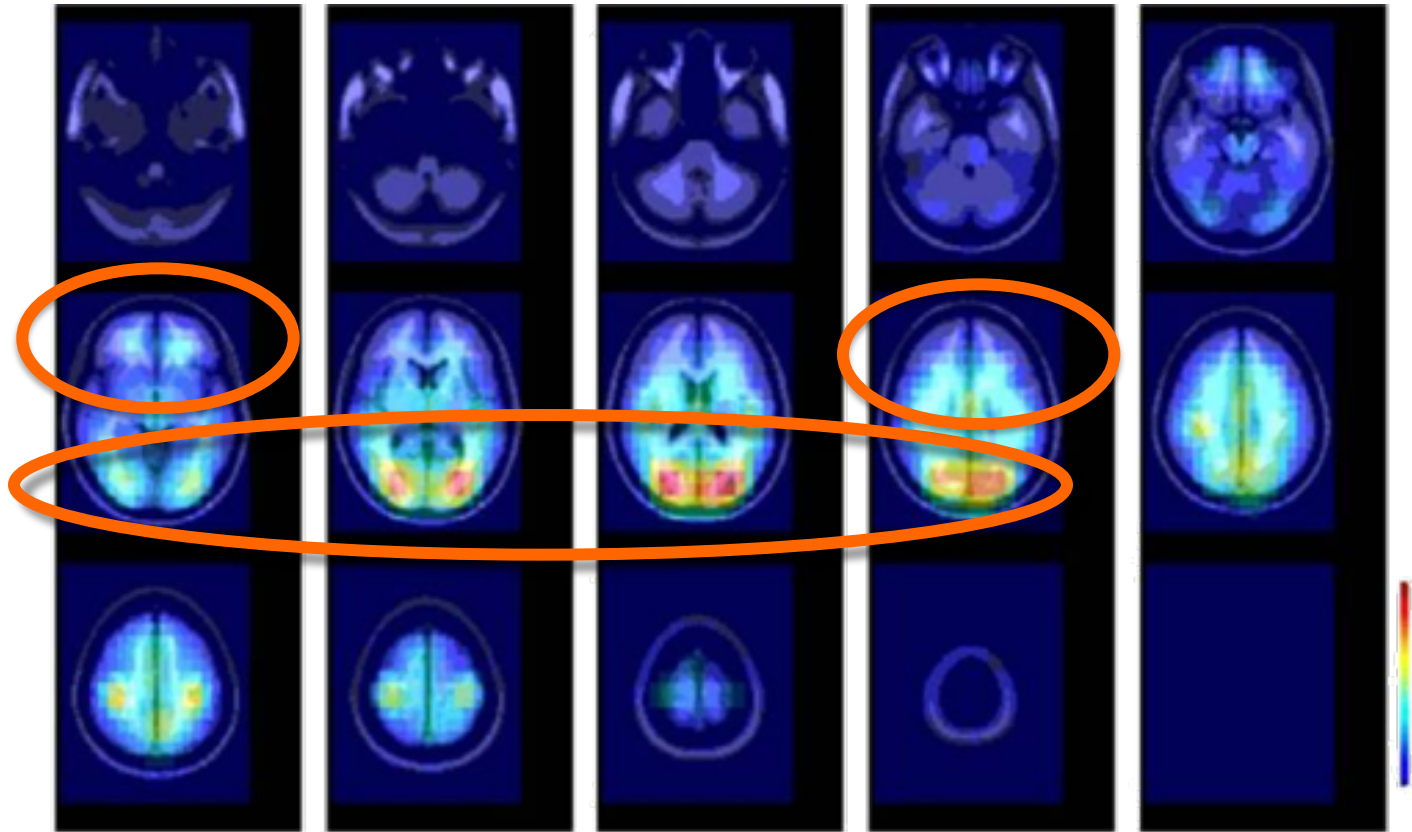
Visual Working Memory



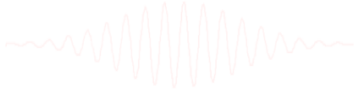
dipoledensity()

Effective Source Density

Eyes-closed emotion imagination

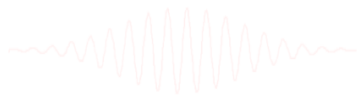
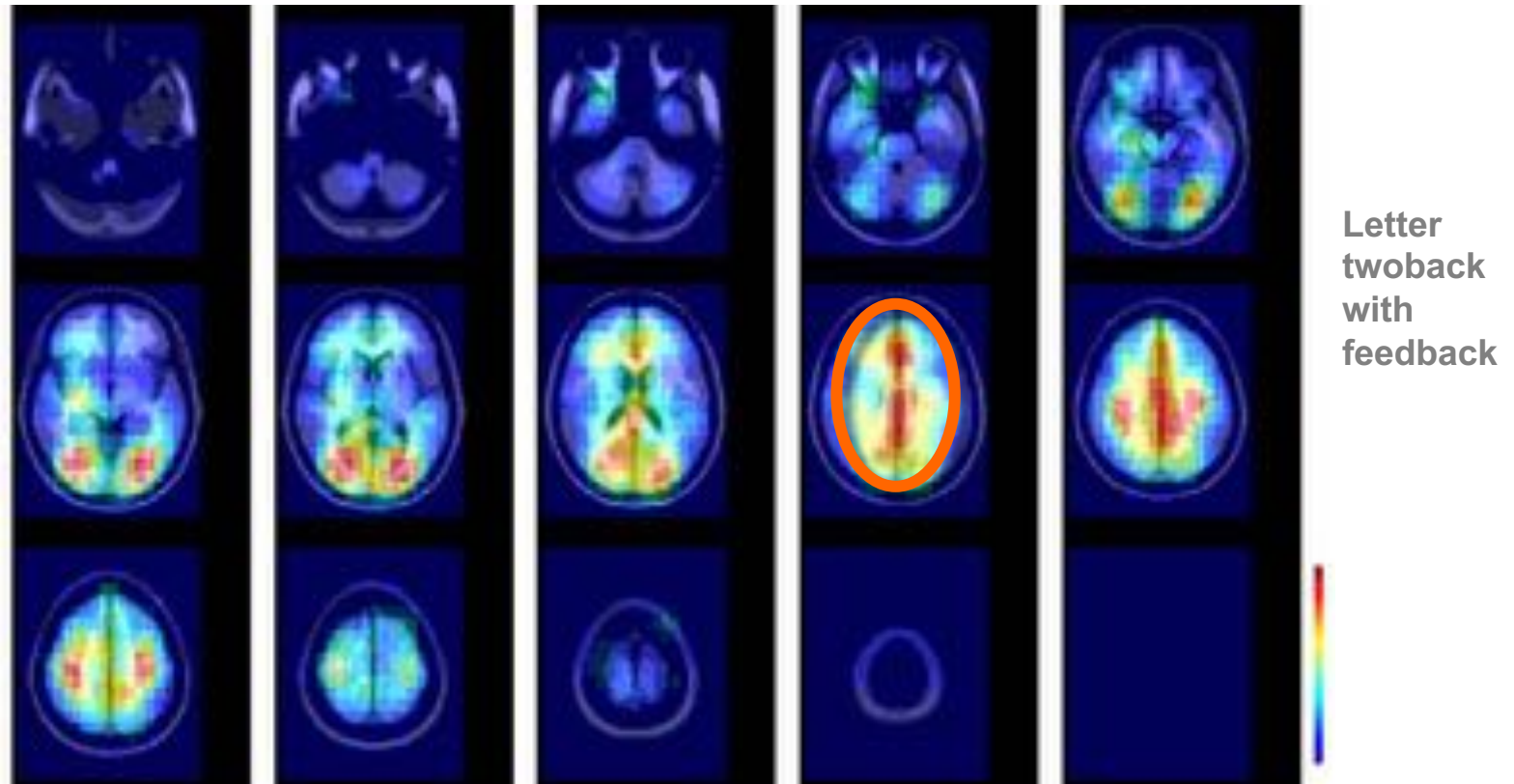


>> dipoledensity()



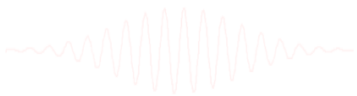
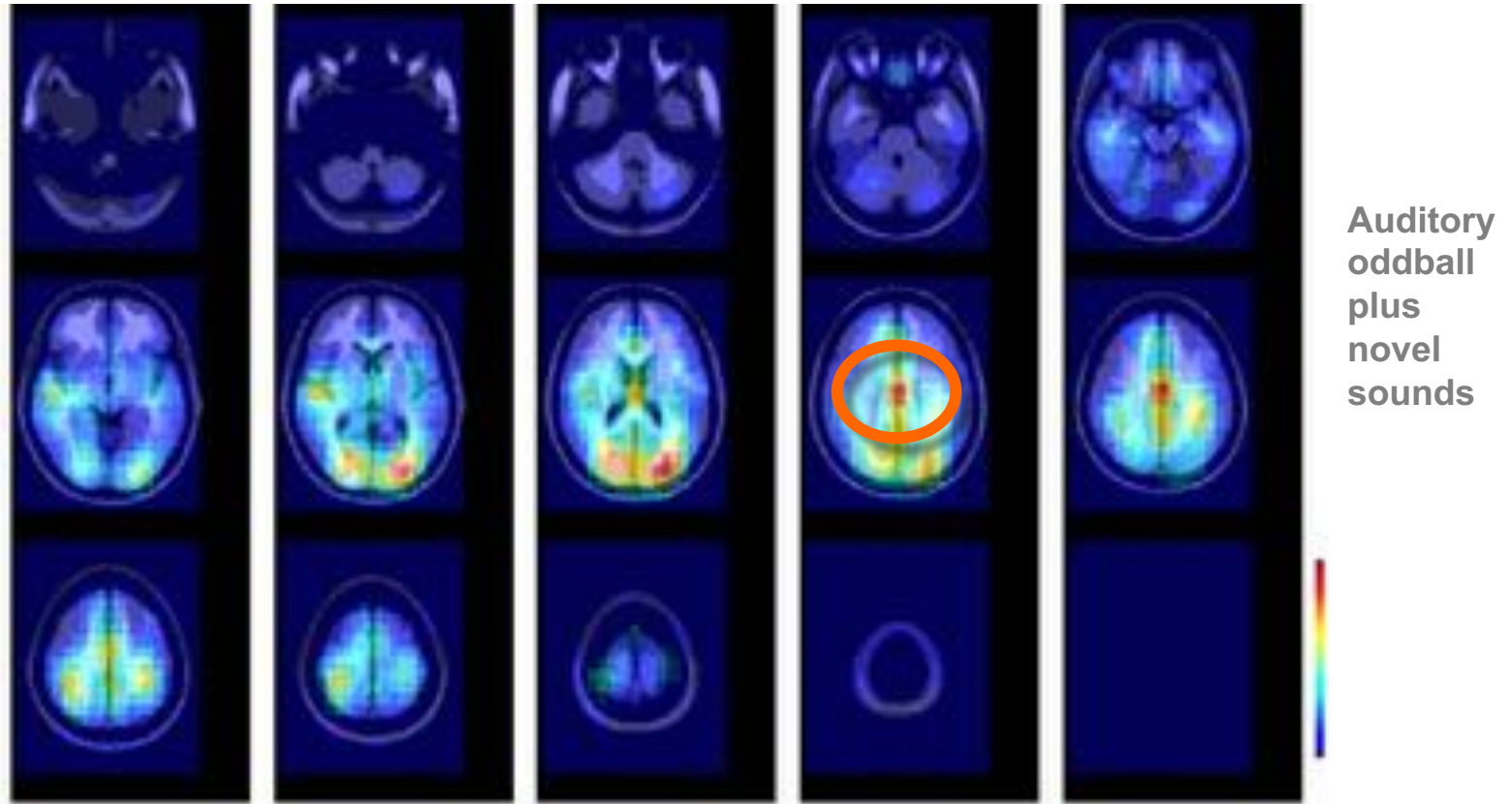
Effective Source Density

Letter twoback with feedback



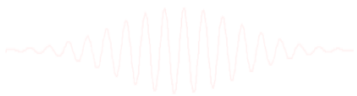
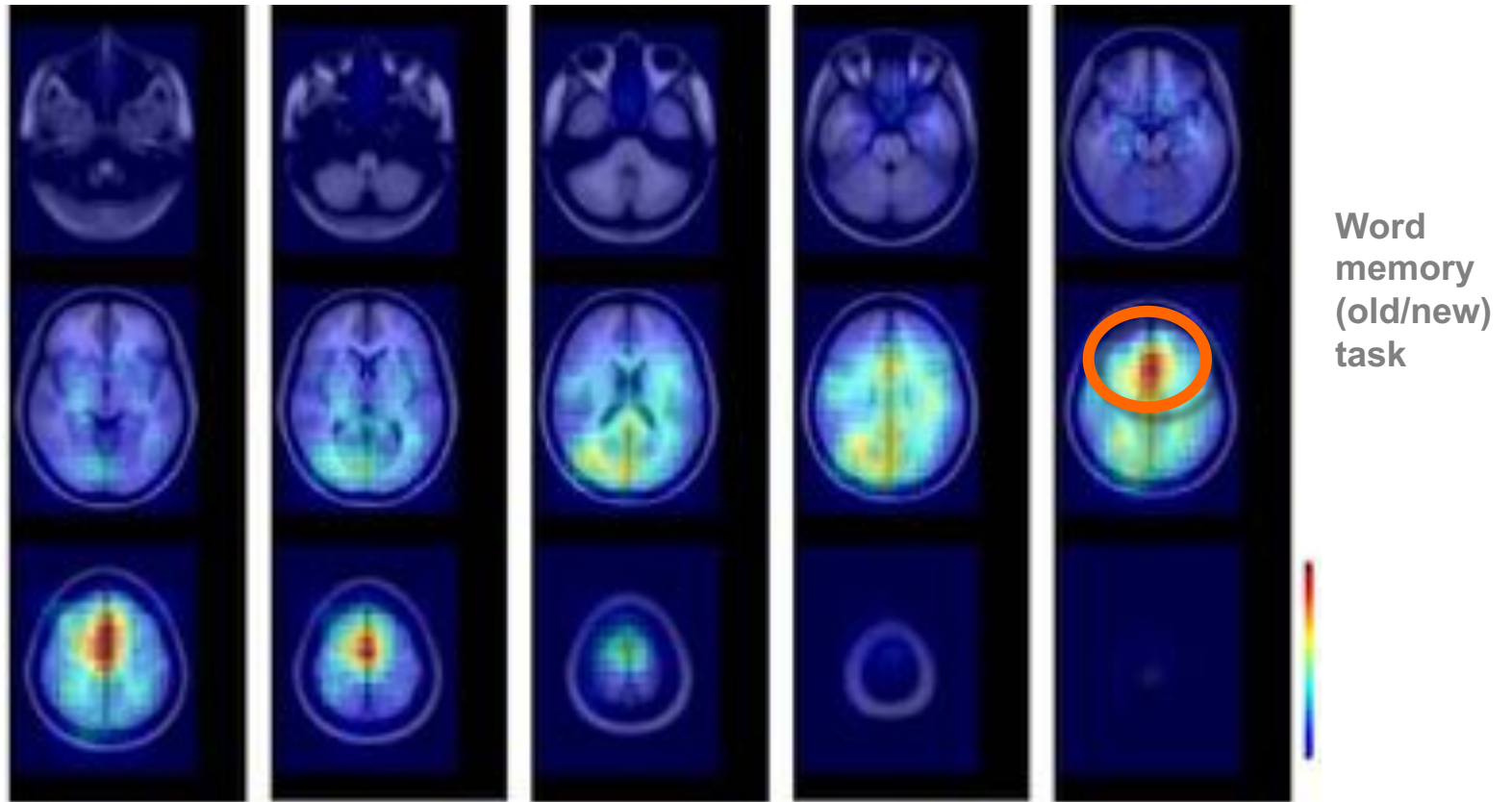
Effective Source Density

Auditory novelty oddball



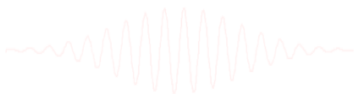
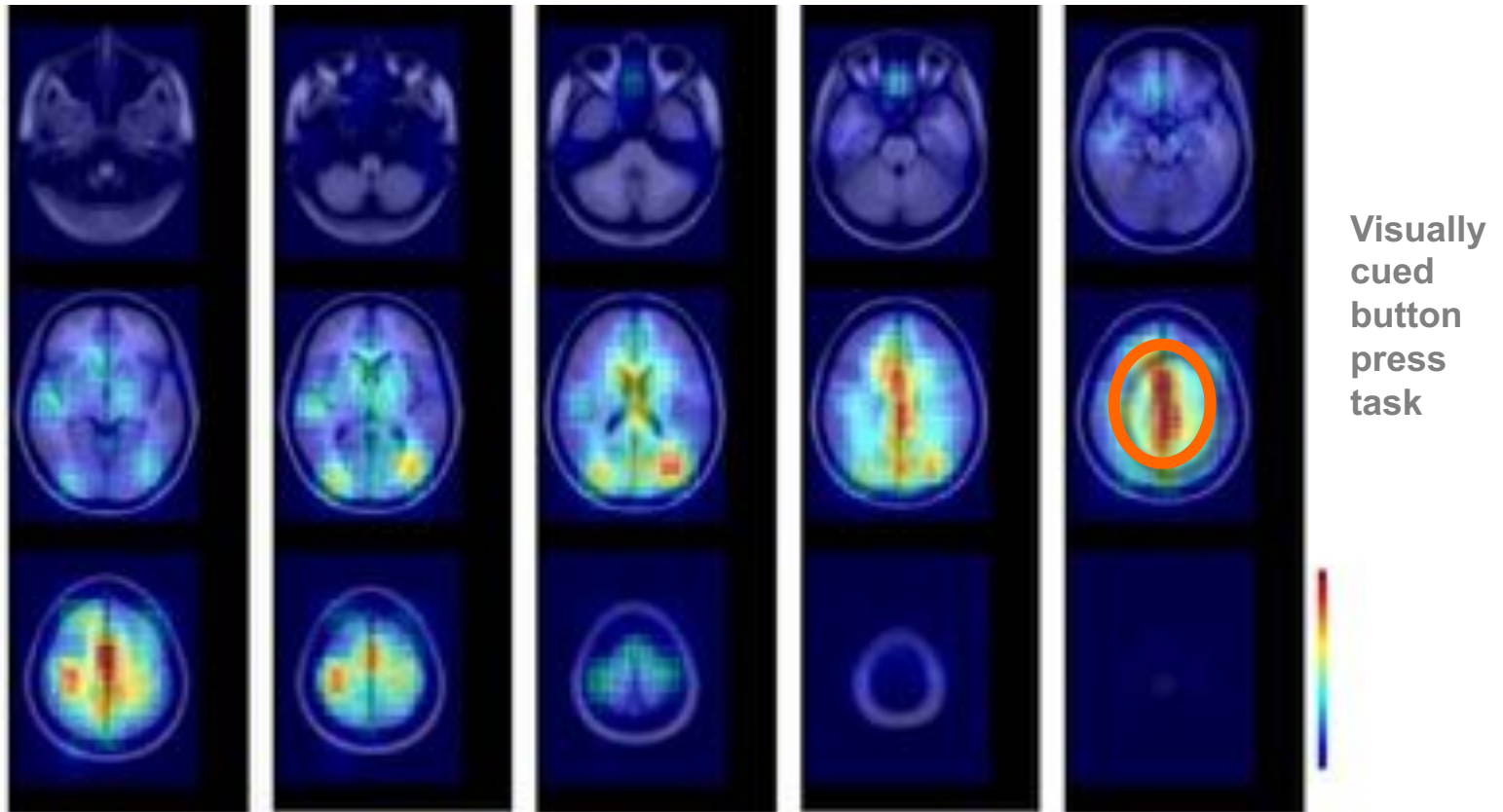
Effective Source Density

A. Old/new word memory

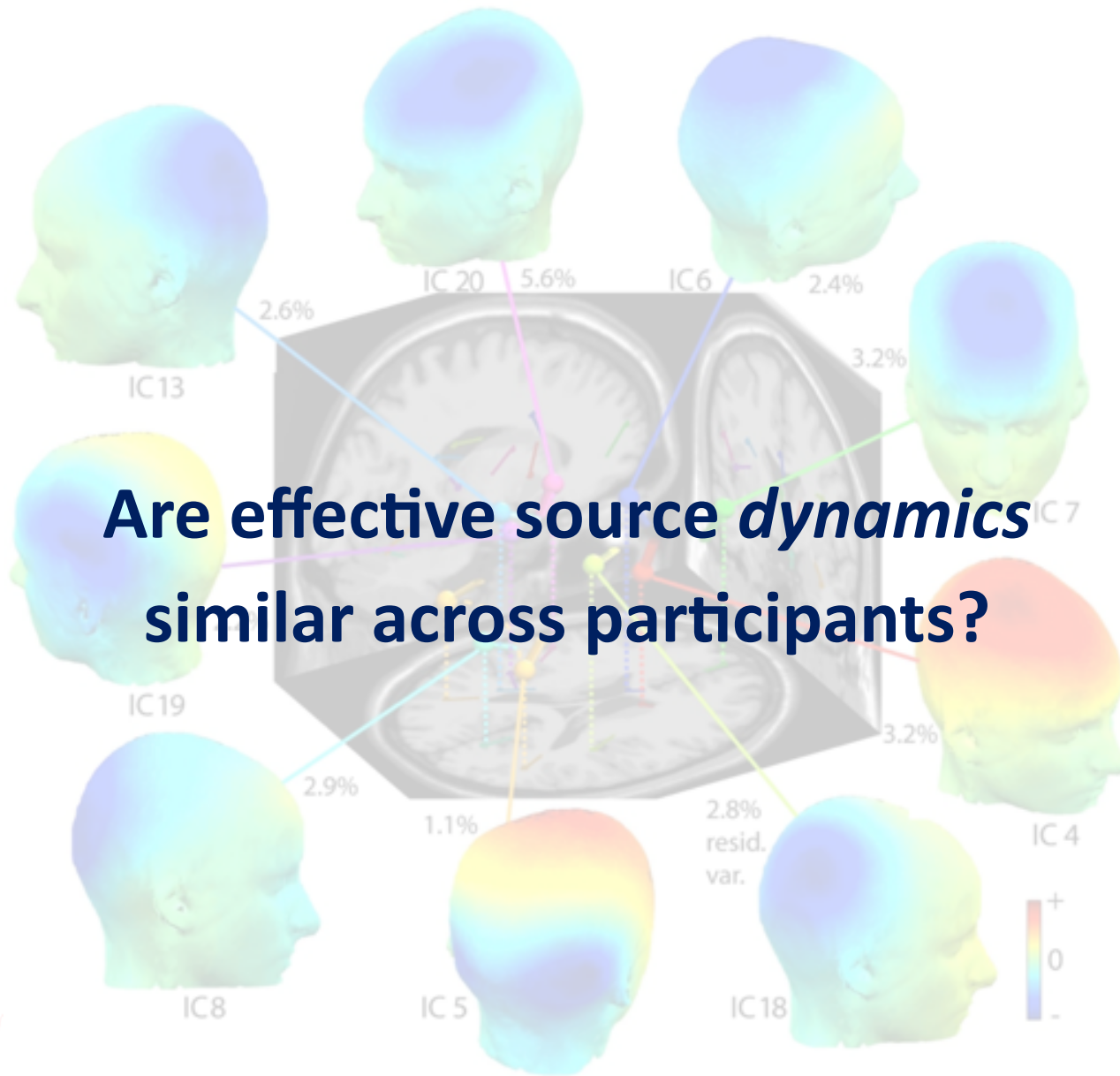


Effective Source Density

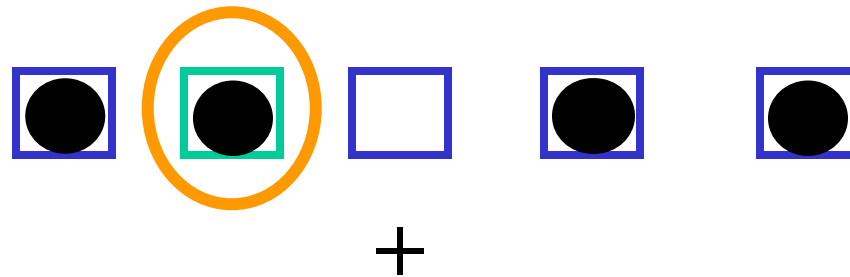
B. Visually cued selective response



Are effective source *dynamics* similar across participants?



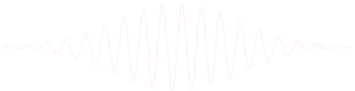
Visual Selective Attention Task



15 subjects

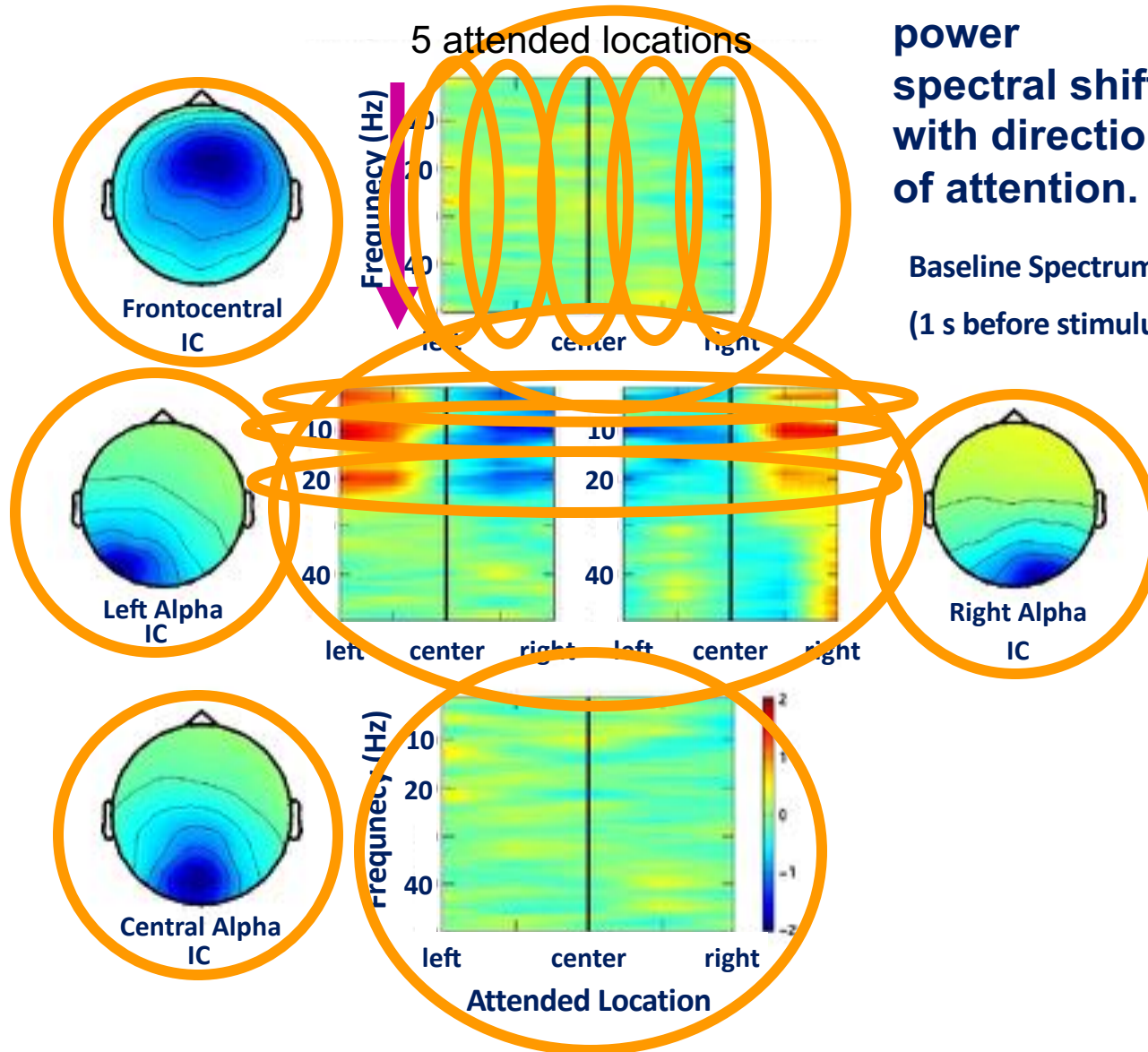
31 channels

Westerfield & Townsend



**Baseline
power
spectral shifts
with direction
of attention.**

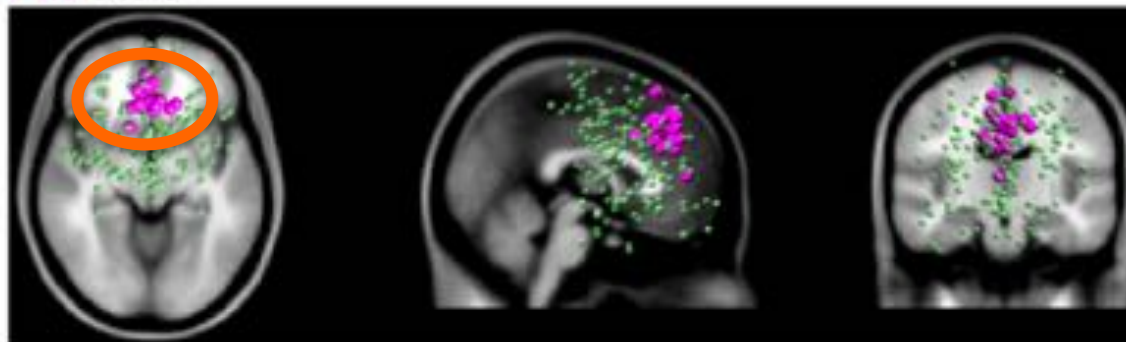
**Baseline Spectrum
(1 s before stimulus)**



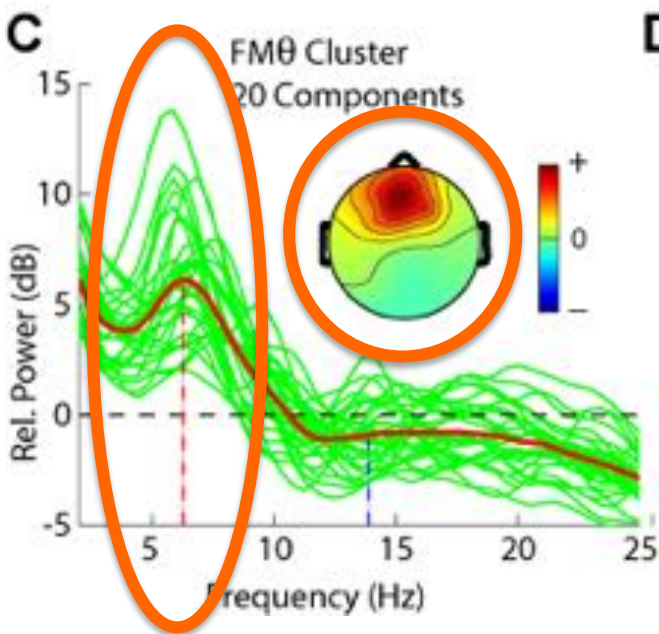
Example: frontal midline theta cluster

B

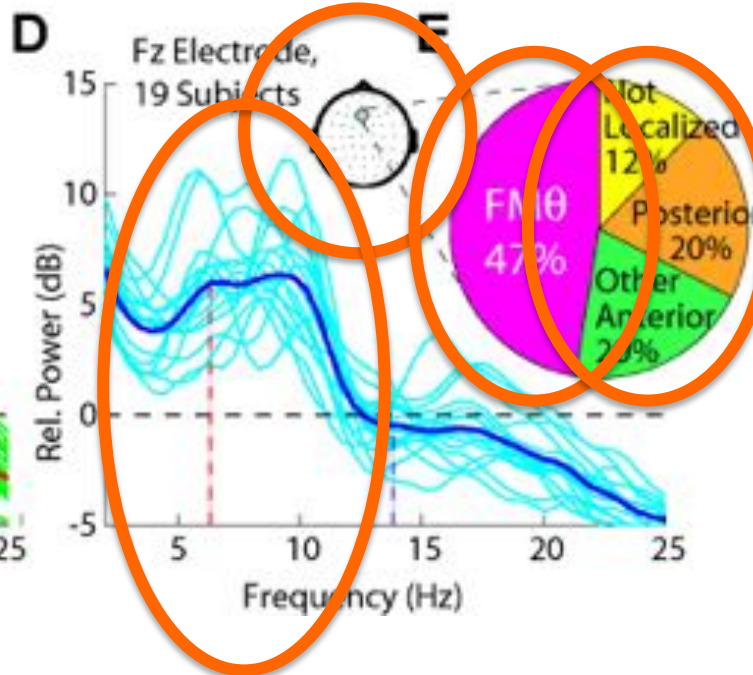
FM θ Cluster



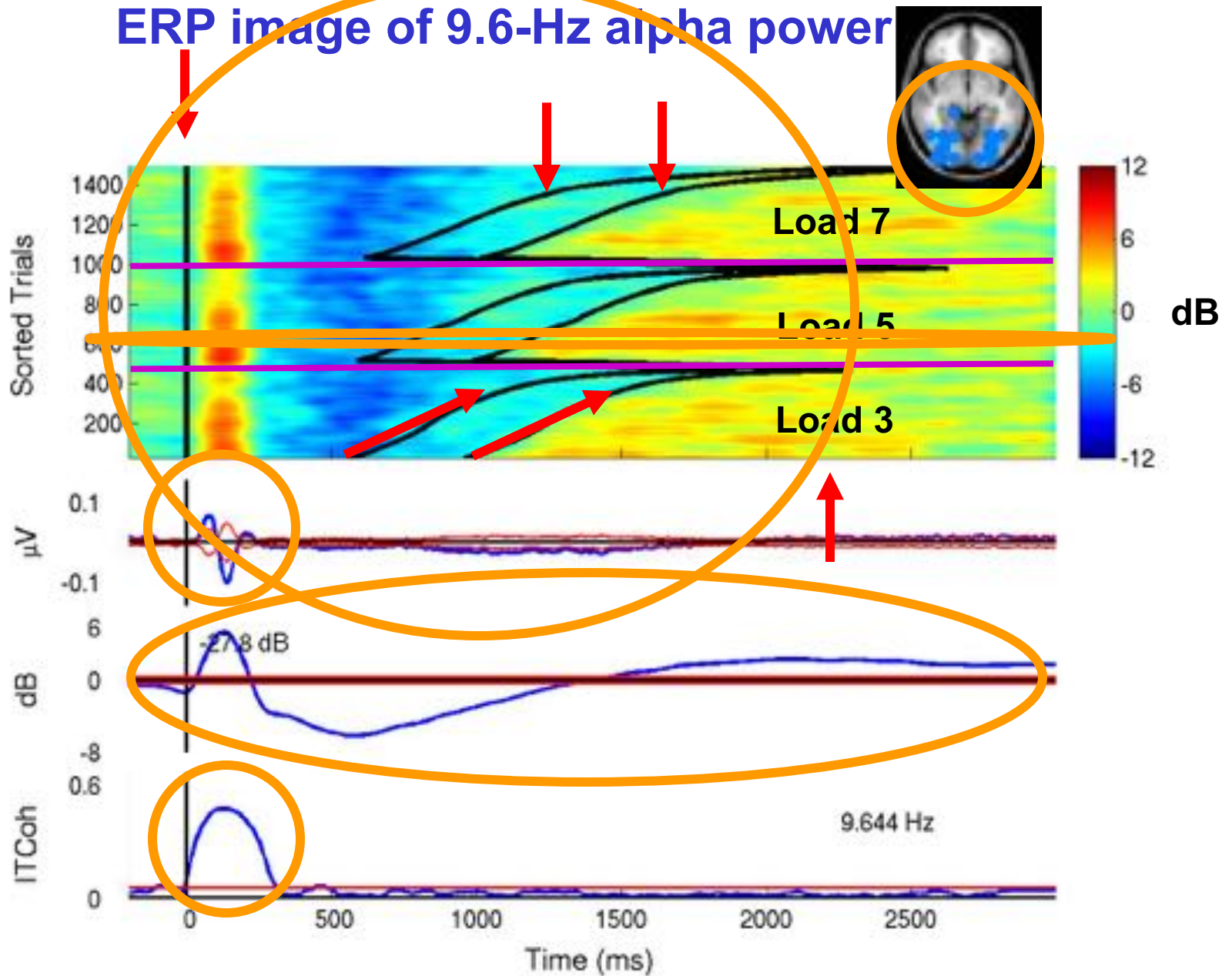
C



D



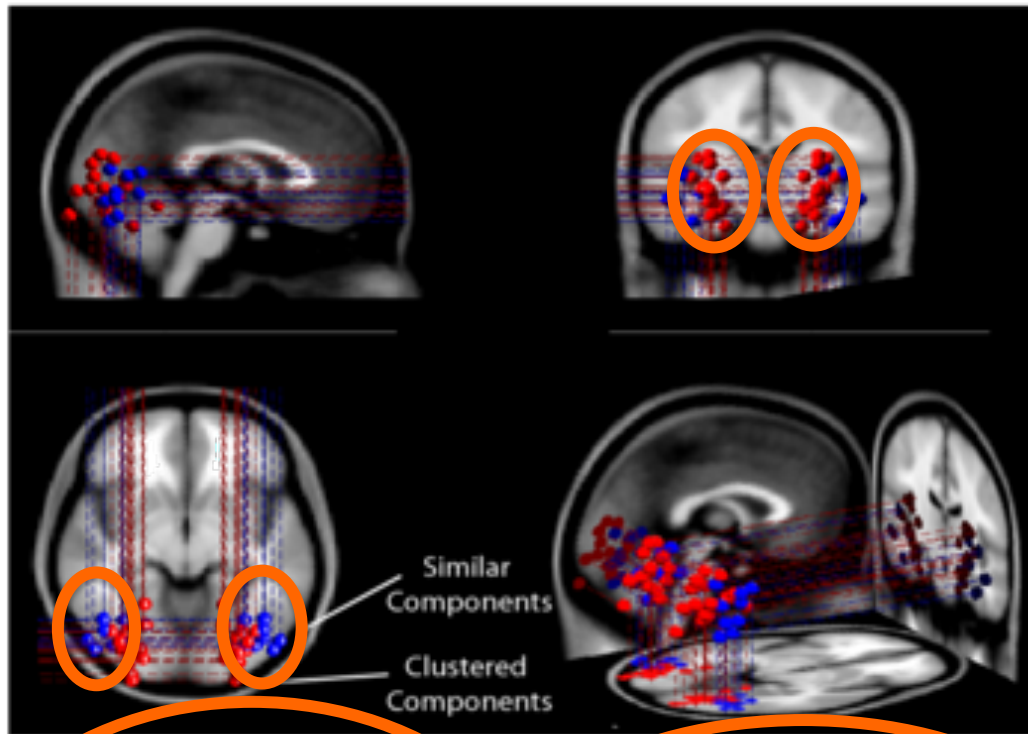
ERP image of 9.6-Hz alpha power



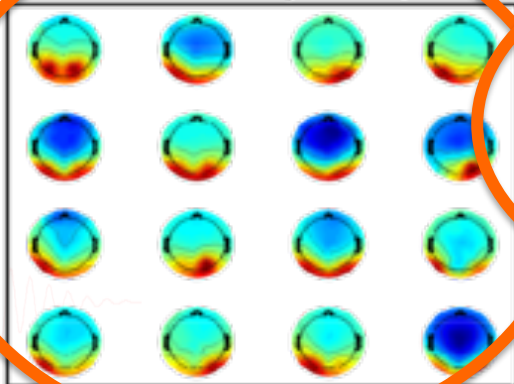
erpimage()

Onton, Delorme & Makeig, 2005.

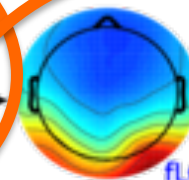
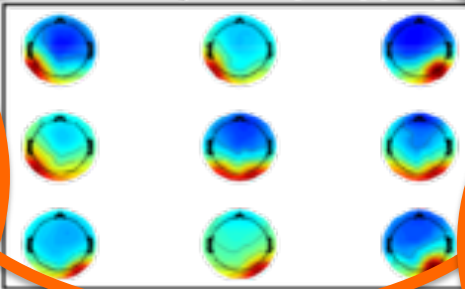
Why don't all subjects contribute to every IC cluster?



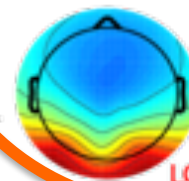
Clustered **LOC** Components (16 Ss)



Similar Components (**FLOC**) (9 Ss)

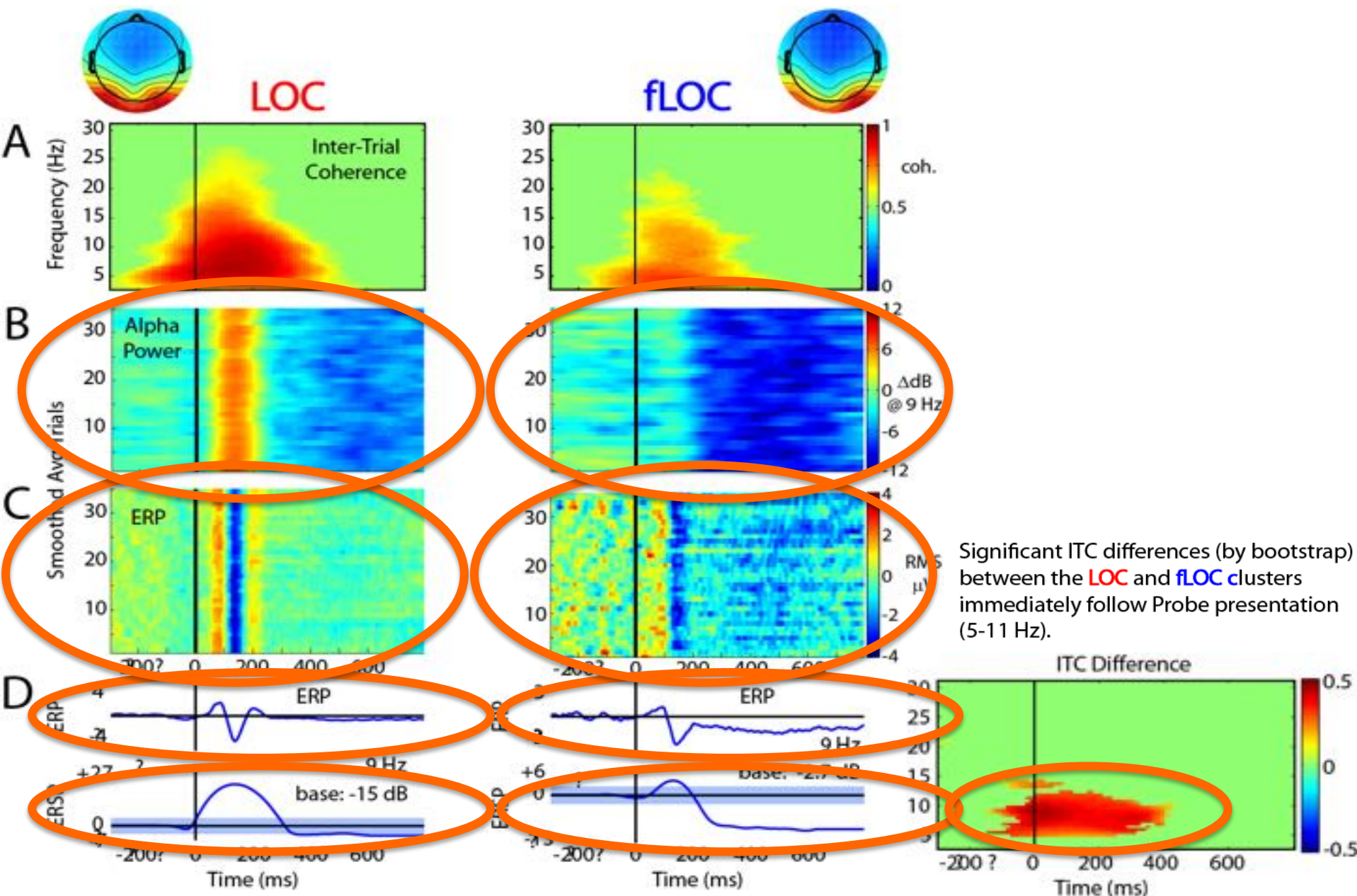


Mean of 9
FLOC Components

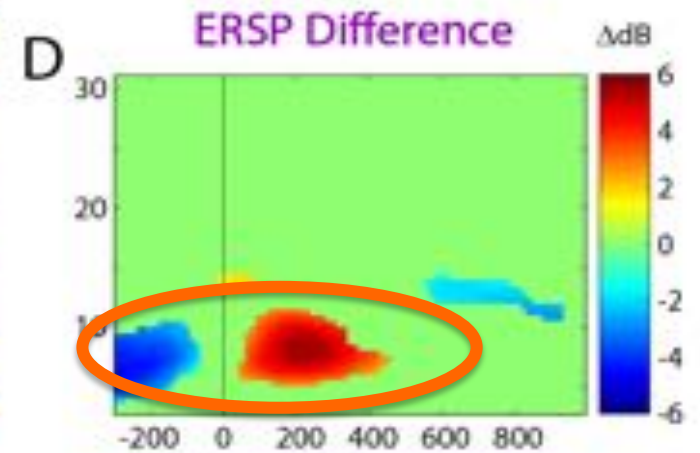
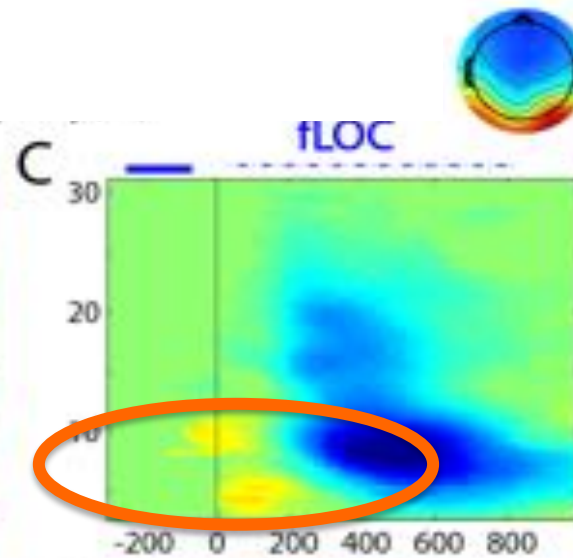
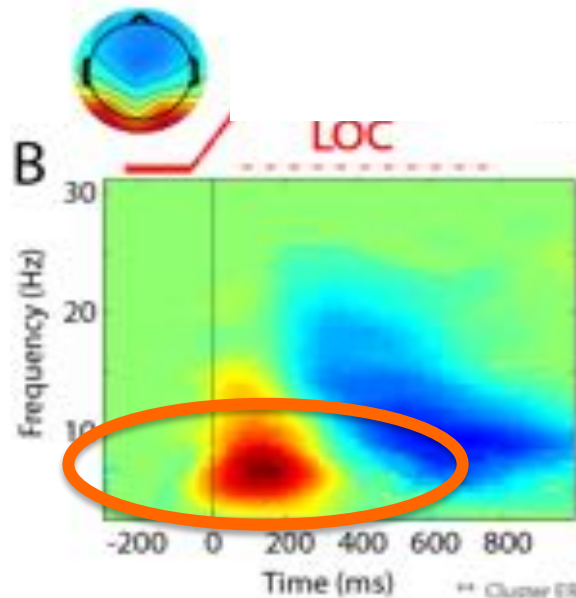
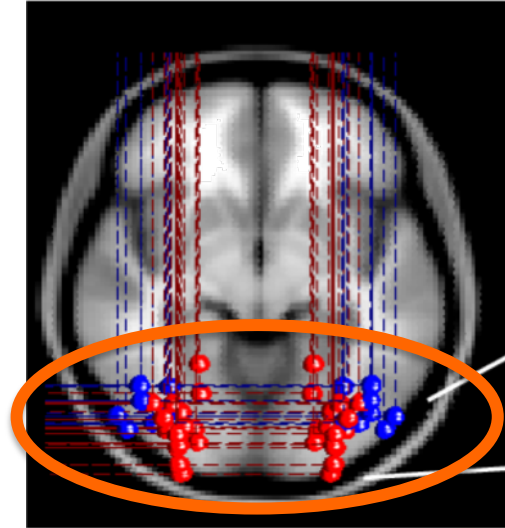


Mean of 16
LOC Components

Subject differences?



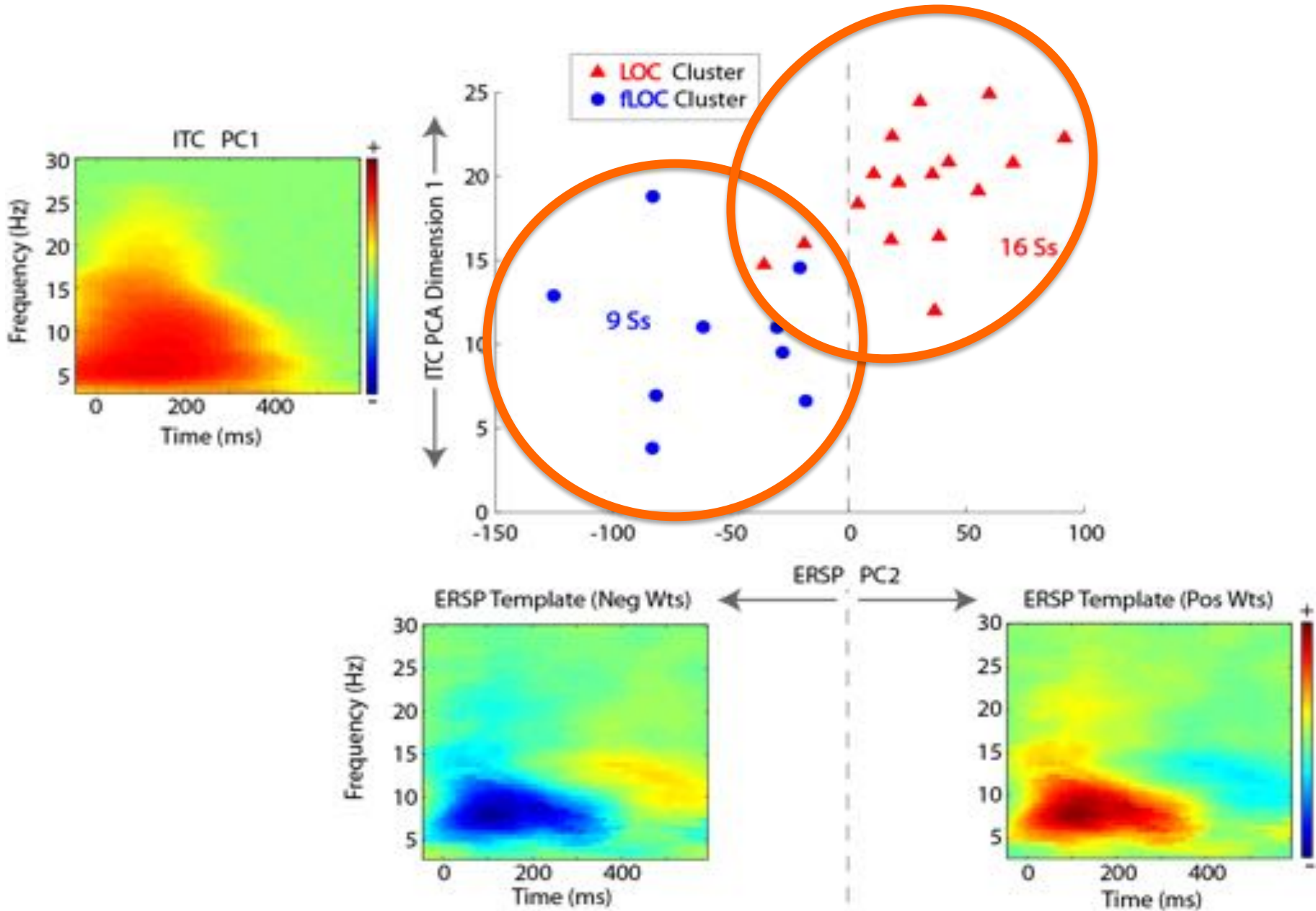
Subject differences?



** Cluster ERSPs show significant activity determined by bootstrap statistics within subject and binomial probability between subjects ($p < 0.01$)

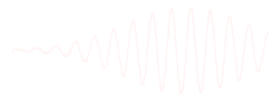
*** Difference ERSP shows significant differences between the two clusters by bootstrap statistics ($p < 0.001$)

Subject differences?



Properties of EEG Independent Components

- Maximally Temporally Independent
- Concurrently Active and Spatially Overlapping
- *Dipolar* Scalp Maps
- **Functionally** Distinct
- Between-Subject Similarity / Complexity



Important Recent Result

Those **linear decompositions** of multi-channel EEG data that find ICs whose time courses are **more** temporally **independent** ...

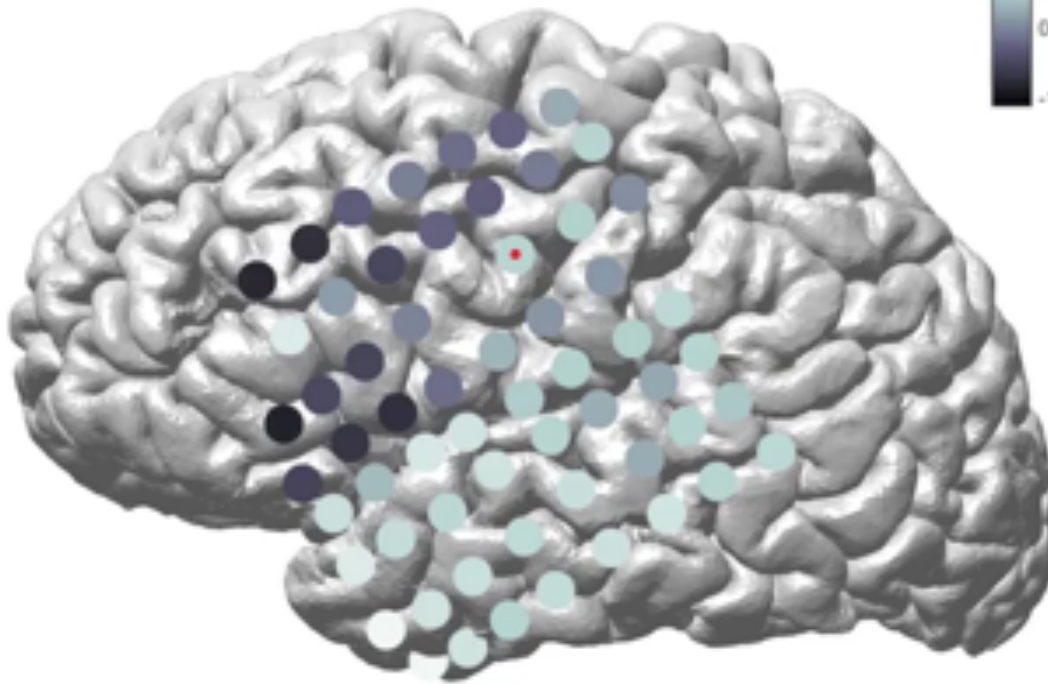
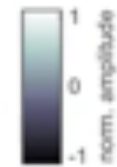
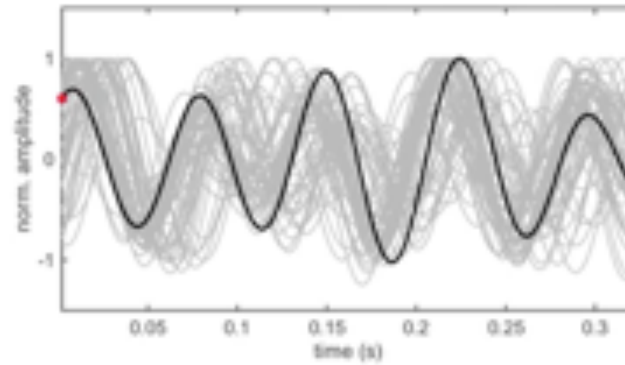
Also find more ICs whose scalp maps are highly '**dipolar**' – i.e., ICs compatible with the spatial projection of a single local cortical effective source (or non-brain artifact) – whose cortical location can be accurately estimated given a good forward-problem head model.

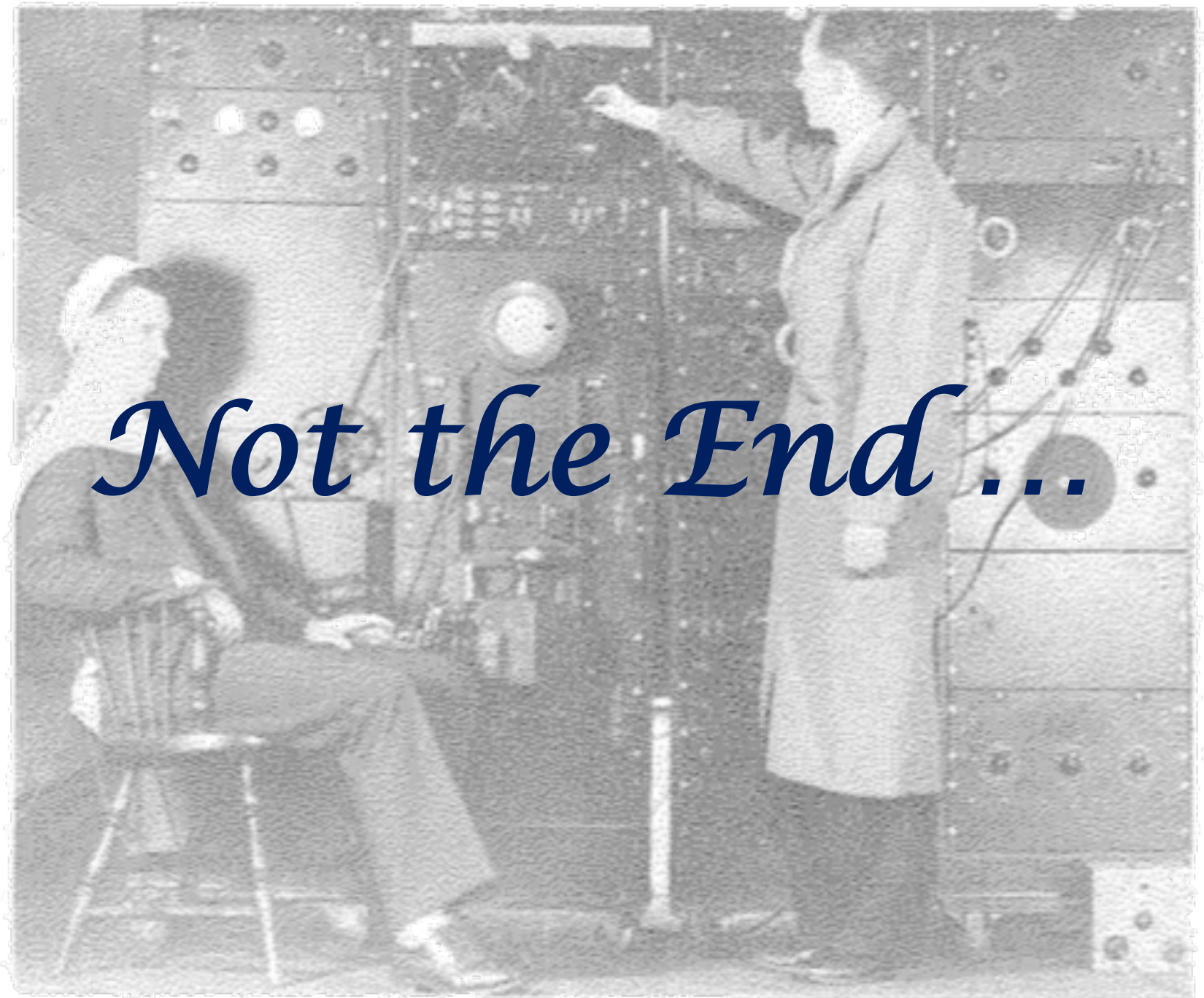
More-independent IC time courses



Larger # of dipolar IC scalp maps

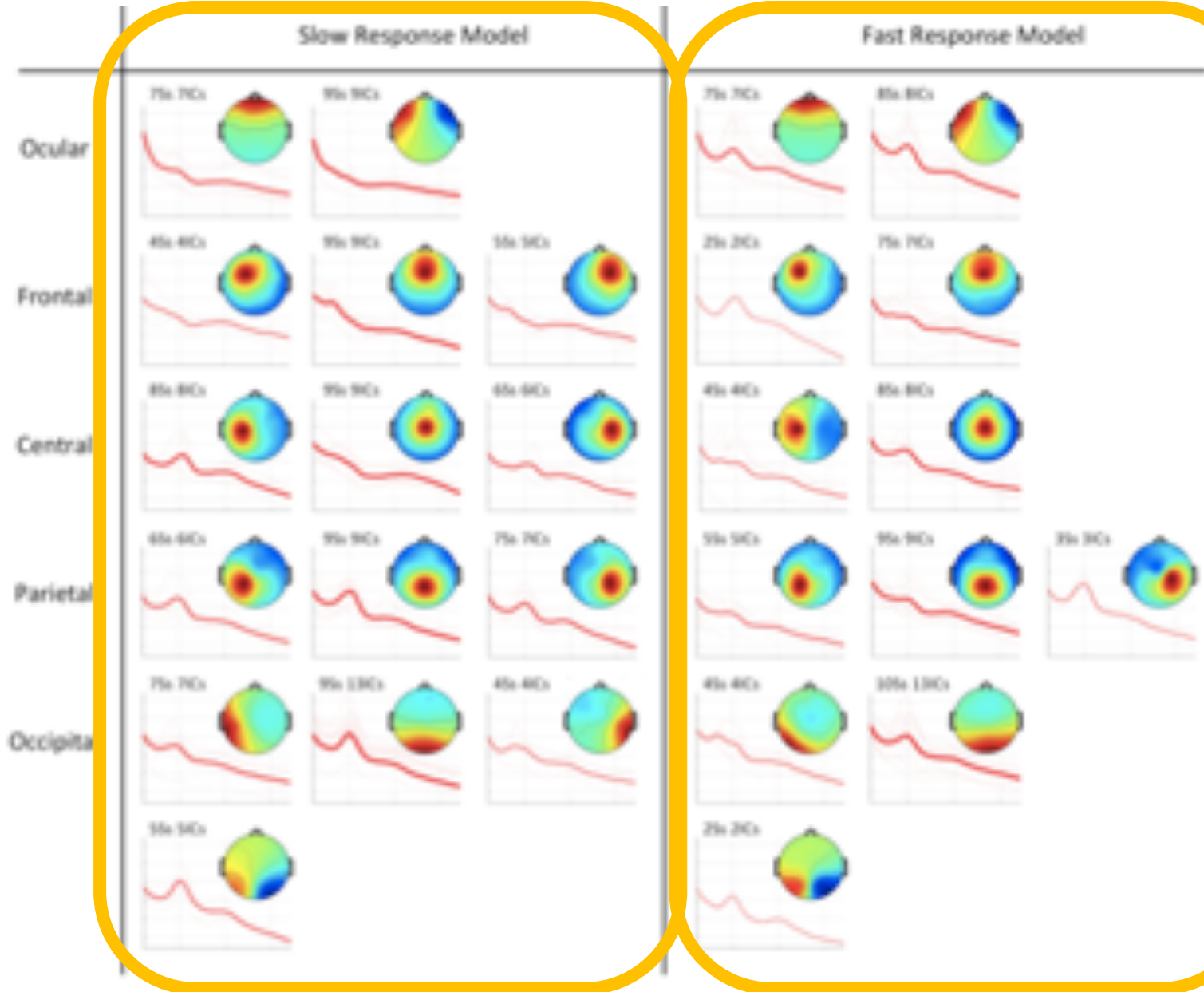
Sleep spindles



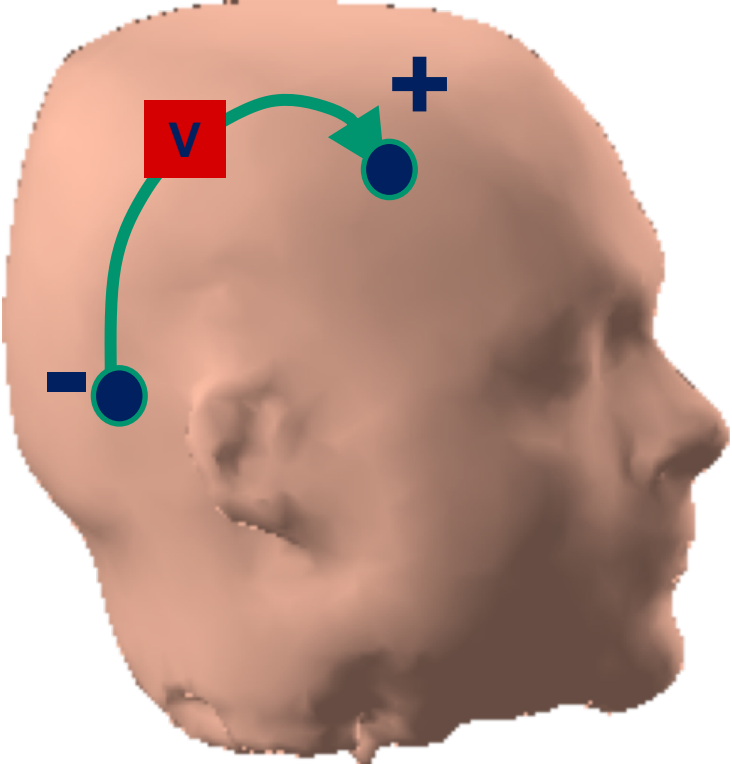


Not the End ...

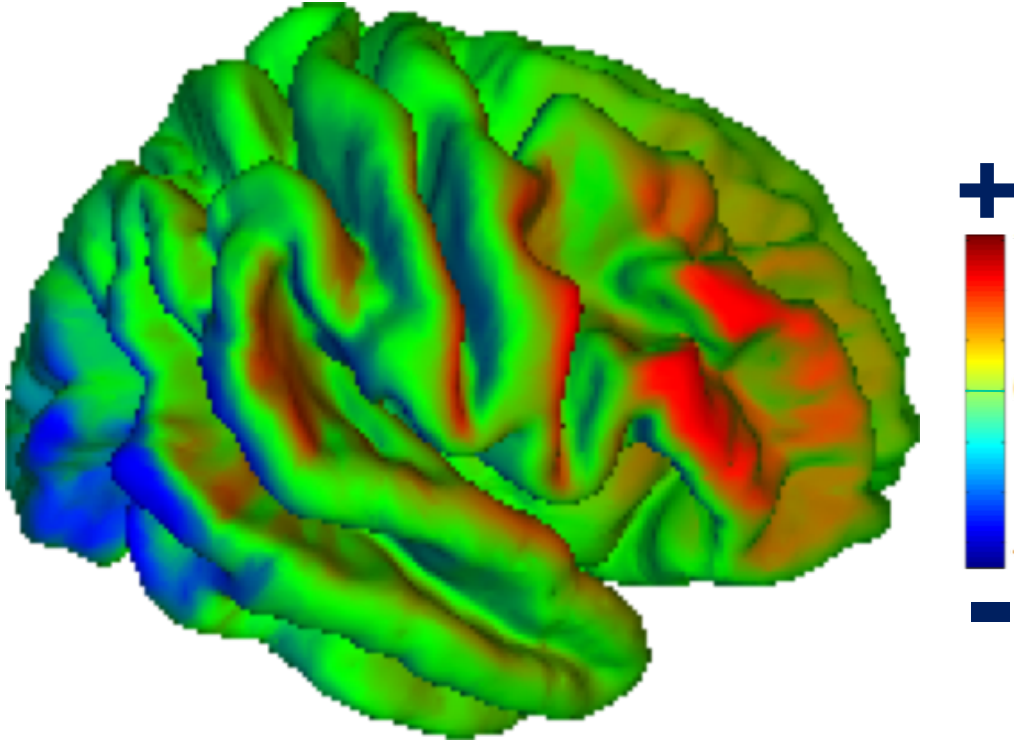
Multi-model ICA: AMICA



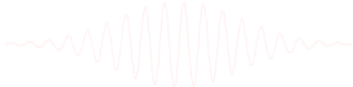
The 'receptive field' of a bipolar EEG channel!



Scalp EEG channel



Its cortical 'receptive field'



Multi-model ICA: AMICA

Reaction speed to simulated driving challenge:

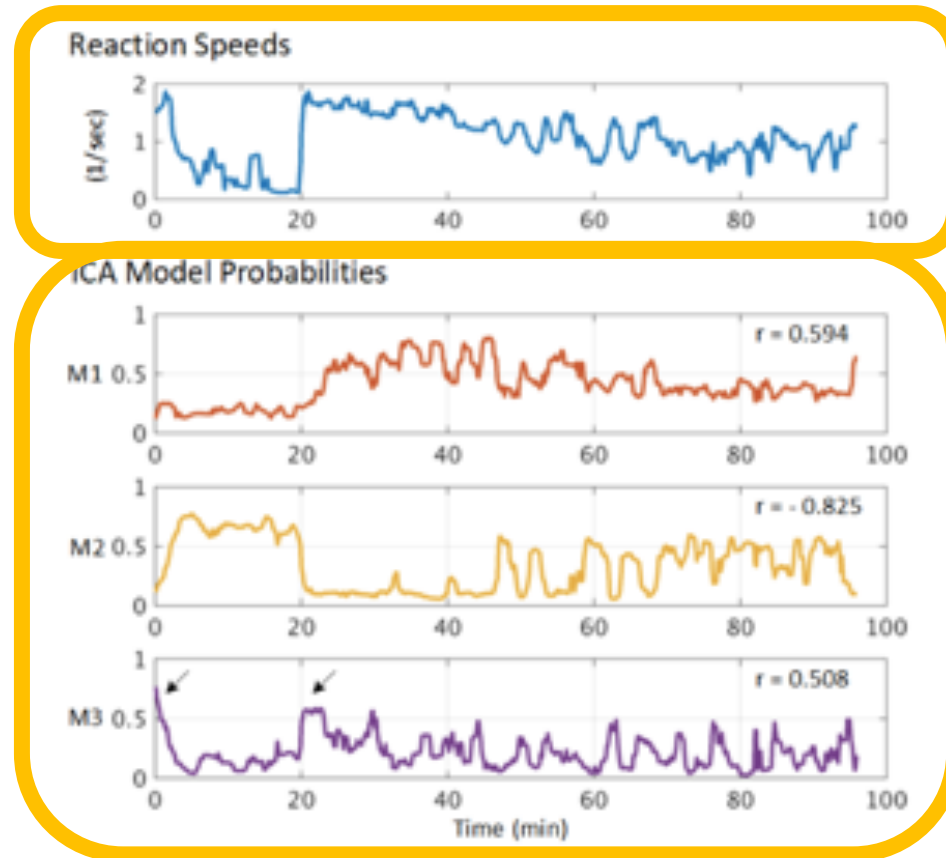


Figure 8: The top panel shows reaction speed changes (inverse of reaction times) in response to lane-departure challenges in one simulated driving session. The three bottom panels show the 5-sec smoothed probabilities of the three ICA models learned by a three-model AMICA decomposition of the whole EEG data session before lane-departure events. Correlation coefficients (r) between each model probability time course and reaction speed are indicated. Black arrows in the lower panel mark brief (alert) periods when model M3 was dominate and reaction speed high.